Information Processing and Management 59 (2022) 103061

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

Federated learning review: Fundamentals, enabling technologies, and future applications

Syreen Banabilah [a], Moayad Aloqaily [b], Eitaa Alsayed [a], Nida Malik [a], Yaser Jararweh [a,*]

[a] *Duquesne University, Pittsburgh, PA, USA*
[b] *Machine Learning Department, Mohamed Bin Zayed University of Artificial Intelligence (MBZUAI), United Arab Emirates*

ARTICLE INFO

Keywords:
Federated learning
Decentralized learning
Distributed learning
Machine learning
Mobile edge networks
Data privacy
Data security

ABSTRACT

Federated Learning (FL) has been foundational in improving the performance of a wide range of applications since it was first introduced by Google. Some of the most prominent and commonly used FL-powered applications are Android's Gboard for predictive text and Google Assistant. FL can be defined as a setting that makes on-device, collaborative Machine Learning possible. A wide range of literature has studied FL technical considerations, frameworks, and limitations with several works presenting a survey of the prominent literature on FL. However, prior surveys have focused on technical considerations and challenges of FL, and there has been a limitation in more recent work that presents a comprehensive overview of the status and future trends of FL in applications and markets. In this survey, we introduce the basic fundamentals of FL, describing its underlying technologies, architectures, system challenges, and privacy-preserving methods. More importantly, the contribution of this work is in scoping a wide variety of FL current applications and future trends in technology and markets today. We present a classification and clustering of literature progress in FL in application to technologies including Artificial Intelligence, Internet of Things, blockchain, Natural Language Processing, autonomous vehicles, and resource allocation, as well as in application to market use cases in domains of Data Science, healthcare, education, and industry. We discuss future open directions and challenges in FL within recommendation engines, autonomous vehicles, IoT, battery management, privacy, fairness, personalization, and the role of FL for governments and public sectors. By presenting a comprehensive review of the status and prospects of FL, this work serves as a reference point for researchers and practitioners to explore FL applications under a wide range of domains.

1. Introduction

Over the past few years, Federated Learning (FL) achieved considerable success in many research areas and different industries. Despite being a newer setup for training Machine Learning (ML) models, FL has been the foundation of improving the performance of applications like Google keyboard (ie. the Gboard on Android) predictive text (Hard et al., 2018) and Google Assistant (McMahan & Thakurta, 2021). FL settings have made it possible to run on-device, collaborative ML (McMahan & Ramage, 2017) that is secure

[*] Corresponding author.
E-mail addresses: banabilahs@duq.edu (S. Banabilah), maloqaily@ieee.org (M. Aloqaily), alsayede@duq.edu (E. Alsayed), malikn@duq.edu (N. Malik), jararwehy@duq.edu (Y. Jararweh).

https://doi.org/10.1016/j.ipm.2022.103061
Received 11 February 2022; Received in revised form 19 June 2022; Accepted 9 August 2022
Available online 26 August 2022

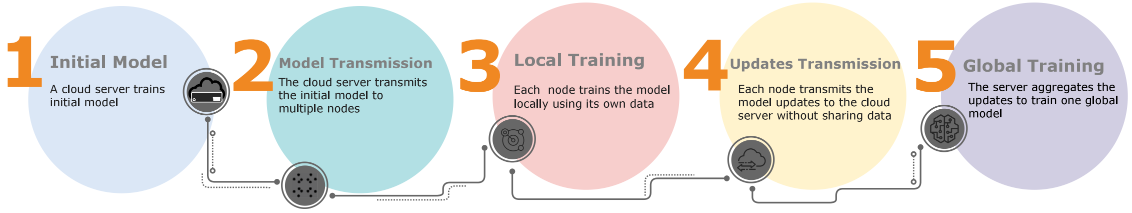

1
Initial Model
A cloud server trains initial model
2
Model Transmission
The cloud server transmits the initial model to multiple nodes
3
Local Training
Each node trains the model locally using its own data
4
Updates Transmission
Each node transmits the model updates to the cloud server without sharing data
5
Global Training
The server aggregates the updates to train one global model

Fig. 1. The training process of a Federated Learning setting.

and private. The collaborative learning happens across multiple clients (or users), where clients' personal data is not transmitted to a central server provider and remain on the local client device.

FL can be described as an on-device, collaborative ML setting. It was introduced by Google in 2015 (Konečṅy, McMahan, Ramage and Richtárik, 2016; Konečný et al., 2016). In FL, as show in Fig. 1, a global ML model that lives on a central service provider is shared across a pool of clients (or users) by being broadcasted to their local edge devices. The global model is used as an initial model and trained on each client's data on the local device, resulting in a trained local model and client data remaining on the client edge device. Once a local model is trained, the updated parameters of the client are sent back to the central server in order to be aggregated across the pool of participating clients and used to update the global model parameters. Once the global model is updated, a training round of FL is complete, and the updated global model is broadcasted to another selected pool of clients to run a second round of training. FL is an iterative and collaborative learning setting. Due to the modularized nature of FL, it fosters an easy environment for private and secure collaborative ML. It is typically at the step of aggregating the client local parameter updates where several secure computations can be incorporated to enforce the privacy and security of client data and the global model trained.

FL contrasts with (1) traditional centralized ML training, where all local client datasets are communicated and shared to a central server, (2) distributed ML training, where parallelization of the training processes is required, and (3) decentralized ML training, where typically it is assumed that local client datasets are independent and identically distributed (IID). Table 1 presents criteria for a complete comparison between centralized ML and FL. Since FL follows a decentralized structure, it trains a ML model locally on edge devices and aggregates local updates from a pool of clients to be communicated to a central server. Whereas the training process in centralized ML occurs on a central server from the start with local client data stored on the central server as well. Also, FL's collaborative learning setting allows for improvements on client model personalization since the local training process uses unique user data to update the global model. Moreover, FL preserves user privacy by not sharing user data and communicating only aggregated model updates that undergo several secure computations in practice.

Most recent surveys on the progress of FL have focused on technical considerations of FL including framework architectures and how they are applied in practice (Yang, Liu, Chen, & Tong, 2019a), enabling software and hardware tools and platforms for FL (Aledhari, Razzak, Parizi, & Saeed, 2020a), system challenges in FL such as communication costs, security and privacy, and resource allocation (Lim et al., 2020). While the majority of prior literature surveys focus on technical considerations and challenges of FL, there has been a limitation in more recent work that presents a comprehensive overview of the status and future trends of FL in applications and markets.

Since FL is massively growing research problem, we present a comprehensive literature of progress in FL and the future of its directions under several areas and disciplines. Our work aims to draw a big picture of the fundamental of FL for those with a basic understanding of ML including clearly defining an FL setting and its components (specifically, ML and Edge Computing), different FL architectures, system challenges within FL, and privacy-preserving methods in FL scenarios. Furthermore, we focus a vast portion of the work on highlighting current and future trends in daily applications and real-world use-cases in industry today that are relying more and more on the FL learning infrastructure.

While previous survey papers focused on the conceptual and technical considerations of FL, the contributions of our survey are:

- An in-depth overview of FL applications and trends in technologies including Artificial Intelligence (AI), Internet of Things (IoT), Blockchain, Natural Language Processing (NLP), Autonomous Systems, and resource allocation.
- An in-depth overview of FL applications and trends in market use cases including domains of Data Science, Healthcare, Education, and Industry.
- The discussion of future open directions and challenges in FL including recommendation engines, autonomous vehicles, IoT, battery management, privacy, fairness, personalization, and the role of FL for governments and public sectors.
- The classification, clustering, and in-depth comparison of a large collection of recent literature progress in FL.

The remainder of this paper is organized as follows. Section 2 discusses the most relevant FL surveys compared to this work. Section 3 presents the background overview, assumptions and characterizations, architectures, and system challenges of a FL setting. Section 4 examines in detail the privacy and security concerns and preservation methods of FL applications in recent works. Section 5 provides a wide scoping of literature of current trends of FL applications in technologies and market industries. Section 6 extends on applications and use cases to highlight the expected and foreseen future trends of FL. Finally, Section 7 concludes the survey.

Table 1
Centralized Machine Learning Vs. Federated Learning.

Criteria	Centralized Learning	Federated Learning
Target	Collect data	Distribute learning
Training	Training on server	Training on edge devices
Aggregation	No aggregation	Aggregation on server
Model	Shared model	Personalized/Shared models
Sharing process	Data sharing	Model updates sharing
Iterations	One-time submission	Iterative process

Table 2
Qualitative comparison of existing Federated Learning surveys.

Survey	Scope	Contributions	Technologies	Year
Lim et al. (2020)	FL background, fundamentals, system challenges, and applications	First FL comprehensive survey about FL features, system challenges, and enabled technology	Deep Learning, Edge Computing	2020
Yang et al. (2019b)	FL definitions, architectures, applications	Design data networks among organizations based on FL	Privacy, Distributed Machine Learning, Edge Computing, Federated Database Systems	2019
Li, Fan et al. (2020)	Explore the main evolution path for existing issues in FL	Propose a detailed review of real-life FL applications in industrial engineering and healthcare	Optimization, Security	2020
Aledhari et al. (2020b)	Background on FL including challenges, advantages, and use cases	Present a thorough deep dive into FL architectures and their use cases	Deep Learning, Machine Learning, Optimization	2020
Ours	FL background, challenges, privacy and security, unique features, characterization, architectures, and real-life applications of FL in technology and markets	First comprehensive survey detailing the technical and application sides of FL	Machine Learning, Deep Learning, Edge Computing, Optimization, Privacy, Security	2022

2. Related work

Despite Federated Learning (FL) being a more recent technology, introduced in 2015 by Google (Konečnỳ, McMahan, Ramage et al., 2016; Konečný, McMahan, Yu et al., 2016), it has shown a significant promise and stride forward. This section presents an overview of the current existing works.

Lim et al. (2020) present a comprehensive survey that shows how Deep Learning on Edge Computing is currently widely used in a variety of applications due to the massive computing power present on edge smartphones today. The authors introduce concepts of FL by discussing the intersection between Deep Learning and Edge Computing. Furthermore, they (Lim et al., 2020) discuss the challenges of the implementation process of running Deep Learning on edge devices with suggested solutions. Finally, their work highlights four existing FL applications within the scope of Edge Computing. In Yang, Liu, Chen, and Tong (2019b), Q. Yang et al. proposes different system challenges of FL with accompanied solutions. They also discuss the three architectures seen in practice today: horizontal FL, vertical FL, and Federated Transfer Learning. Each architecture is discussed from a technical stand point and presented with current applications. At the end, the authors prove that Federated Learning is an efficient technique to build data networks among multiple organizations while preserving user privacy. Li, Fan, Tse and Lin (2020) discuss the development process of FL and how the common challenges faced are resolved during implementation. Furthermore, the author discuss the future direction of FL applications, while primarily focusing on the domain of industrial engineering. The work of Aledhari, Razzak, Parizi, and Saeed (2020b) presents differences between FL and traditional, centralized Machine Learning models. The survey covers a background of enabling technologies that are used to learn a model under a FL setting.

Table 2 presents a summary comparison of existing surveys in the domain. The summary discusses each surveys scope of discussion, unique contributions, technologies explored, and the year of publication. Currently, there is limited literature that reviews the current status of Fl and more particularly in terms of current and future trends of FL applications in technology and markets. Our work aims to fill this gap and serves as a reference point for researchers and practitioners to explore FL applications under a wide range of domains.

3. Federated Learning fundamentals

3.1. Background

In this section, we provide a background overview where we present the two underlying technologies of Federated Learning (FL) – Machine Learning and Edge Computing – as shown in Fig. 2, and formally define FL.

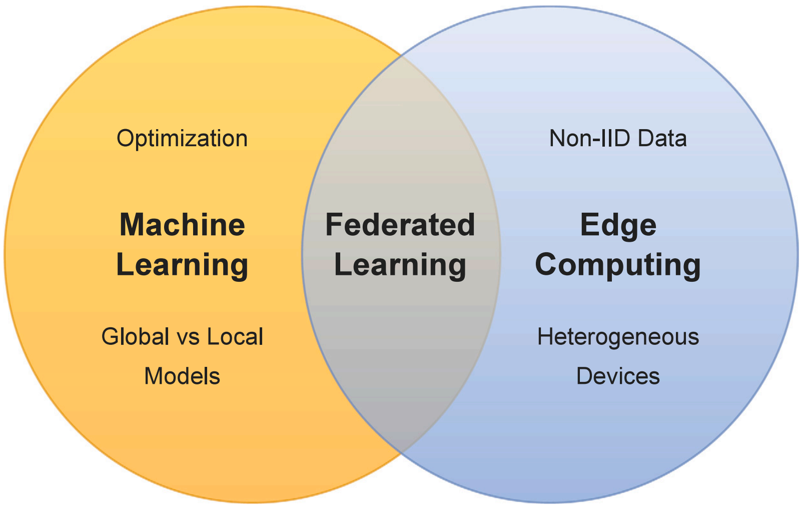


Fig. 2. The orchestration of different components and characteristics that contribute to the Federated Learning setting.

3.1.1. *Machine Learning models*

Over the past two decades, Machine Learning (ML) has become a primary source of information and pattern extraction from large data sources to automate and predict tasks across endless domains. ML algorithms have gone as far as outperforming humans on tasks such as defeating the chess champion, Garry Kasparovimage (Knight, 2020), and winning against the world's number one Go player, Ke Jie (BBC News, 2017). Trained ML models are able to automate tasks such as classifying images (Huang, Liu, Van Der Maaten, & Weinberger, 2017), navigating self-driving cars (Bojarski et al., 2016), recognizing and understanding human speech (Chan, Jaitly, Le, & Vinyals, 2015), and diagnosing diseases (Elayan, Aloqaily, & Guizani, 2021b).

ML refers to mathematical models that are derived using a training dataset of N samples $D = \{X, Y\}$, where $X = \{x_1, x_2, .., x_N\}$ are input samples and $Y = \{y_1, y_2, .., y_N\}$ are labels that map each input sample to a true desired output for the given problem (Bishop & Nasrabadi, 2006). The training dataset is used to extract patterns and define a parameterized or non-parameterized model representation, that is, the learned ML model. ML is a very wide domain, and it inherits concepts from many other related fields such as computer science, statistics, and optimization.

ML algorithms can typically be broken down into three primary attributes:

- A hypothesis function that represents the mathematical model to be learned.
- A cost function that defines the mathematical objective the algorithm aims to achieve.
- An optimization algorithm that solves the cost function to reach the target objective.

ML can be divided into several learning paradigms, some of the most prominent being: supervised learning, unsupervised learning, and semi-supervised learning.

Supervised learning refers to a learning paradigm in ML where the model is learned using a *labeled training dataset* to extract the relationship between the input samples X and the true labels Y (Pahwa & Agarwal, 2019). The purpose of supervised learning is to correctly predict the label for a new, unseen data sample. Ideally, the training dataset distribution would closely mimic real world data samples, and thus, the model is able to leverage the pattern learned from the training dataset samples to map into the unseen world. Supervised learning includes applications such as classifying spam emails as spam or not spam (Renuka, Hamsapriya, Chakkaravarthi, & Surya, 2011) and predicting the future price of Bitcoin (McNally, Roche, & Caton, 2018).

Unsupervised learning is a learning paradigm in ML where the model is learned using an *unlabeled training dataset* to find similar characteristics across the input samples X (Hänsch & Hellwich, 2009). Unlike supervised learning, there are no desired output values since the label is not provided. The purpose of unsupervised learning is to identify hidden patterns across data samples. Examples of unsupervised learning include applications such as clustering DNA patterns for evolutionary biology analysis (James, Luczak, & Girgis, 2018) and predicting buyer purchase patterns (Singh, Mittal, & Pareek, 2016).

Semi-supervised learning is a hybrid learning paradigm where the input dataset is partially labeled, and the unlabeled data samples largely exceed the amount of labeled samples (Lan, Deng, & Chen, 2011). Both supervised learning and unsupervised learning are used during the training process. Unsupervised is generally first applied to learn the pattern structure of the unlabeled dataset and cluster samples, while supervised learning is used to predict the label of the unlabeled samples by learning the pattern within the pool of labeled training samples. The purpose of this type of learning is to improve the performance of ML by using unlabeled data, which is typically not expensive to collect (Lan et al., 2011). Semi-supervised learning can be applied to a wide range of applications with some examples being classification of large corpora of text documents (Al-Laith, Shahbaz, Alaskar, & Rehmat, 2021) and analysis of image or audio data (Gururani & Lerch, 2021; Sohn et al., 2020).

Table 3
Cloud Computing versus Edge Computing.

Aspect	Cloud Computing	Edge Computing
Speed	Data processing that is not time urgent.	Real-time data processing.
Connectivity	Requires reliable Internet connection.	Accessible in areas with limited Internet connectivity.
Communication	Applicable to dynamic data workloads.	Applicable to large data workloads.
Data	Insensitive data stored on central server.	Private data stored on local devices.

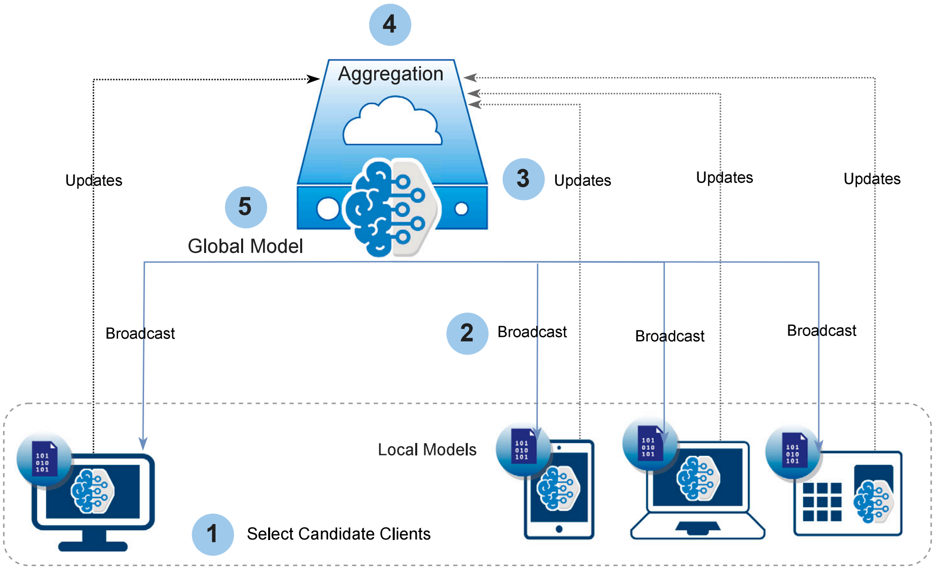


Fig. 3. A single training round in a Federated Learning setting.

Deep Learning (DL) is a branch of ML that uses a hierarchical learning architecture that mimics the human neural brain architecture in order to extract complex, non-linear patterns in data. Similar to ML, DL can apply any of the aforementioned learning paradigms to extract hidden patterns and structure in data. The research community has seen a large focus on DL methods since the introduction of these models in 2006 (Bengio, Courville, & Vincent, 2013). It is with the advancements in DL that we have seen models be able to reach comparable, and occasionally higher, performance to humans on tasks such as image classification (Dodge & Karam, 2017).

3.1.2. Edge Computing

Cloud Computing was a transcending paradigm that brought forth on-demand computer resources such as storage and computing power. The availability of Cloud Computing brought forth a number of benefits including: cost cuts, performance, speed, reliability, and security. The introduction of the Internet of Things (IoT) challenged Cloud Computing by exposing its limitations in handling and communicating enormous amounts of data where challenges of latency, privacy, and security arise. These challenges brought forth the introduction of computing processes that take place at the edge of the network, commonly termed as Edge Computing (Noble et al., 1997).

Edge Computing is a newer computing paradigm that helps distribute computing processes and resources while bringing them closer to the data sources, thus, improving privacy and security concerns (Wang, Zhang, Wang, Ma and Liu, 2020). The most prominent advantage of Edge Computing is cost reduction while maintaining maximum operational efficiency. This reduction in cost with high efficiency is due to the fact that the data is processed at the edge device, eliminating the need to move it to a central service provider and allocate large storage (Ghosh & Grolinger, 2021). Furthermore, since computations are happening at the edge, on or near the edge devices, computations are faster than through Cloud Computing, presenting an added advantage especially for where real-time response is required. Table 3 shows a comparison between Cloud and Edge Computing.

To complete painting the picture of the advancements in computing infrastructures, Edge Computing presented limitations where several edge IoT applications compete for resources. To resolve competition for resources on limited edge devices, fog computing was introduced. Fog computing serves a similar purpose to Edge Computing but introduces a higher level of abstraction, where computations are run on devices in the Local Area Network (LAN) to which the multiple IoT or edge devices are connected. Fog computing overcomes the challenge of limited resources for competing workloads by running computations of a small-scale cloud that lives at the network edge, referred to as a cloudlet.

3.1.3. *Federated Learning definition*

Federated Learning (FL) can be defined as a ML setting where different clients (or users) collaborate to learn a model on a central server while keeping client data decentralized (Kairouz et al., 2021).

In traditional ML settings, the model resides on a central server. Data from clients is transferred to the central server and used to train a global model using all client data. This global model is a general model that captures patterns or information that is general to the pool of clients as a whole. The global model is then shipped to the client edge devices to be used for inference.

Two primary limitations arise under the traditional setting. The first limitation is on privacy and security concerns in transferring sensitive client data to a central service provider. The second limitation is in the global model representation being general and not representing personalized patterns of each unique client.

FL offers a shift in traditional ML settings that fosters a hybrid training setup that reaps the benefits of a central global model and a pool of decentralized local models. FL resolves the limitation of privacy and security by allowing client data to remain on the local edge devices and overcomes the limitation of a global, general representation by training local models and aggregating their contribution into a global central model.

Starting with a global model on a central server, a training round (or iteration) in FL can be described as follows:

1. A selection process identifies a pool of candidate clients that will participate in the FL round.
2. The global model is broadcasted to the selected client edge devices.
3. Each client computes a local update on the model using the local data. The update is, for example, computed using Gradient Descent.
4. The central server collects and computes an aggregated model update from the local clients. The aggregation step can aggregate several processes, which may include a secure aggregation technique for privacy preservation or lossy compression to enhance communication efficiency (Kairouz et al., 2021).
5. The central server updates the global model representation by the aggregated update computed from the local clients.

Finally, after several training rounds, the global model performance is tested before deployment to end clients for inference use (Wei et al., 2020). Fig. 3 represents a single FL training round.

3.2. *Assumptions and characterization of Federated Learning*

Federated Learning (FL) settings are defined by set of characteristics and assumptions depicted in Fig. 4. This section formalizes the definition of FL by exploring assumptions and traits of user data, edge devices, optimization and training considerations, and privacy preservation.

3.2.1. *Non-IID data*

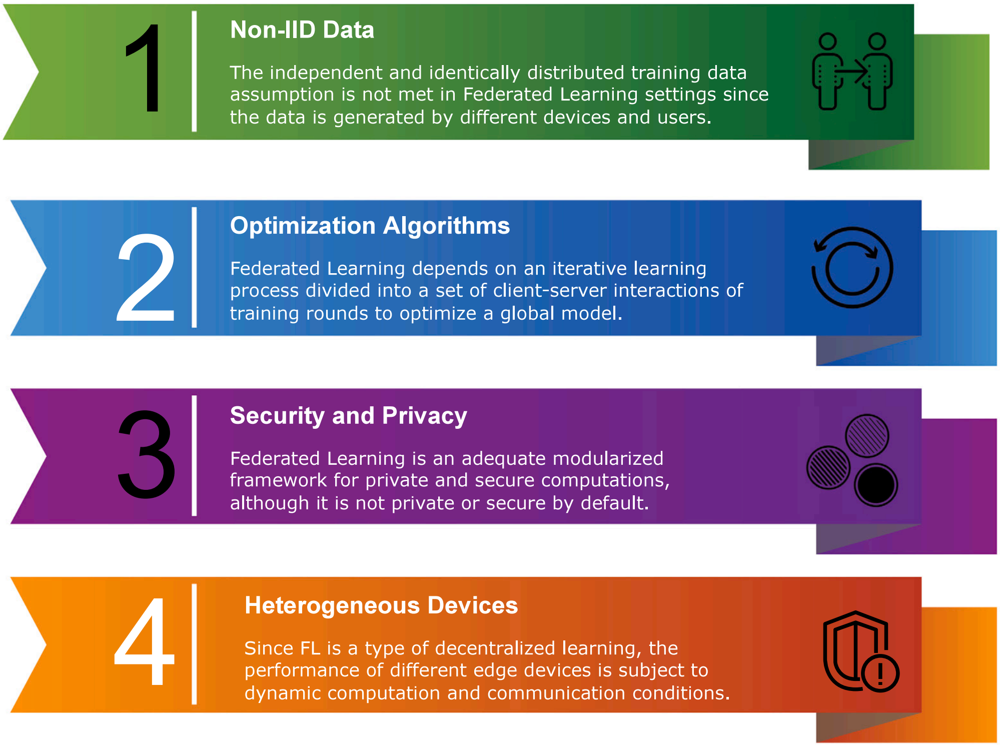
The independent and identically distributed (IID) assumption about training data is often made in standard Machine Learning (ML). In ML algorithms where training is centralized, the data is collected from different sites and stored in a single shared storage. However, FL trains a global model across data generated from distributed devices, which participate in an FL round across different unique users, time zones, and geographical locations. Therefore, the IID assumption is violated, and FL ends up heavily dealing with non-IID client data.

Consider the following: FL selects a client $i \sim \mathcal{Q}$, and then during local training, samples are drawn for each client following $(x, y) \sim \mathcal{P}_i(x, y)$. We can characterize the non-IID nature of the client data in relation to two aspects. First, the more common nature of the non-IID data stems from the fact that each client is typically a unique end user. This brings forth an effect similar to the data drift effect seen in traditional ML, where there exists a distribution shift between the train and test datasets. This distribution shift can be thought of as the difference between $\mathcal{P}_i$ and $\mathcal{P}_j$ sampled from two unique clients, i and j respectively. For example, shifts between $\mathcal{P}_i$ and $\mathcal{P}_j$ may be due to a skew in feature distribution, skew in label distribution, and/or imbalanced quantity of data samples across clients. Second, another distribution shift to consider in a FL setting is the change in $\mathcal{Q}$ (i.e. participating clients) and $\mathcal{P}_i$ (i.e. client behavior) as time passes (Kairouz et al., 2021).

3.2.2. *Optimization algorithms*

The FL training process takes place across ‘rounds’, which are an iterative process divided into a set of client–server interactions aimed at achieving a high performing global model. Each round of this process begins with broadcasting the global model to the participating client edge devices, training local models on each unique client to compute a new local model update at each edge node, and finally, aggregating the local client updates and imposing the aggregated update to the global model at the server end.

In traditional ML, Gradient Descent (GD) is the standard optimization algorithm used with the goal of minimizing the gradient of the cost function with respect to the model parameters being optimized. However, GD requires connections of low-latency and high-throughput. This constraint makes GD inadequate for training ML models in a non-centralized fashion, such as in the scenario of FL where we require collaborative learning across edge client devices. The author of Wang et al. (2019) showed that using GD to optimize a ML model under a FL setting did minimize the cost function, however, under a limited resource budget. The work showed a convenient trade-off between local updates and the global parameter aggregation. However, because FL deals with a massive amount of data across distributed clients, using the same optimization algorithm as in centralized ML will delay the learning process and significantly impact the convergence speed of FL. Furthermore, according to Tran, Bao, Zomaya, Nguyen, and Hong (2019), data size, power constraints, and wireless channels all have a negative impact on standard GD optimization in FL



Non-IID Data

The independent and identically distributed training data assumption is not met in Federated Learning settings since the data is generated by different devices and users.

Optimization Algorithms

Federated Learning depends on an iterative learning process divided into a set of client-server interactions of training rounds to optimize a global model.

Security and Privacy

Federated Learning is an adequate modularized framework for private and secure computations, although it is not private or secure by default.

Heterogeneous Devices

Since FL is a type of decentralized learning, the performance of different edge devices is subject to dynamic computation and communication conditions.

Fig. 4. A summary of the set of features and assumptions that characterize a standard Federated Learning setting.

settings. These negative effects can be described under two folds. The first one is the performance of learning accuracy, which is constrained by the computational and communication resources of the unique clients, and the other appears in the constraint on convergence time of FL training which demands energy consumption on both the edge client and the wireless network.

The aforementioned challenges make it necessary to select an appropriate optimization algorithm for a FL setting. The authors in Nilsson, Smith, Ulm, Gustavsson, and Jirstrand (2018) compare three FL optimization algorithms and their performance including: Federated Averaging (FedAvg), Federated Stochastic Variance Reduced Gradient, and CO-OP. The results showed that FedAvg performed the best. While FedAvg actually works well under non-IID data assumptions, it is optimized for homogeneous network settings. The authors in the article (Cai, Lin, Zhang, & Yu, 0000) display a dynamic sample selection optimization algorithm, FedSS, in order to deal with heterogeneous networks in FL settings. To improve the performance of models trained in FL settings, the authors of Chen et al. (2019) present an algorithm that optimizes for both joint learning and wireless resource allocation.

3.2.3. *Security and privacy*

In itself, Federated Learning (FL) is not considered a secure or private ML setting. However, FL provides an adequate setting in which the ML process is modularized into independent units and maintains data on local client devices. As a result, FL does provide an added layer of privacy to the clients by keeping local data separate from the central service provider. Nevertheless, clients and the central server in a FL setting are still susceptible to security breaches. Some examples of security breaches or privacy concerns include information leakage of raw client data, attacks that prevent the global server model from being updated, or attacks that to bias the global model to generate inferences that favor a particular entity or party (Kairouz et al., 2021).

Aside from training the global model through local client updates, we describe four common secure computations that support in minimizing security and privacy concerns in FL settings (Yang et al., 2019b).

- Secure multi-party computation is an area of cryptography that allows parties to collaborate on running a predefined computation using all party inputs but where each party will only access a subset of the outputs. This method may suffer from limitations when dealing with real numbers, since cryptography methods generally are equipped to deal with finite data (Yao, 1986).
- Differential privacy involves augmenting client data with noise to hide original information about clients before being shared. This method balances a trade-off between privacy and accuracy of the local model updates that are communicated to the central server (Dwork, 2008).
- Homomorphic encryption is a type of encryption method that allows to run mathematical operations on encrypted data. This means that there is less of a chance of receiving any information about the raw data in comparison to a differential

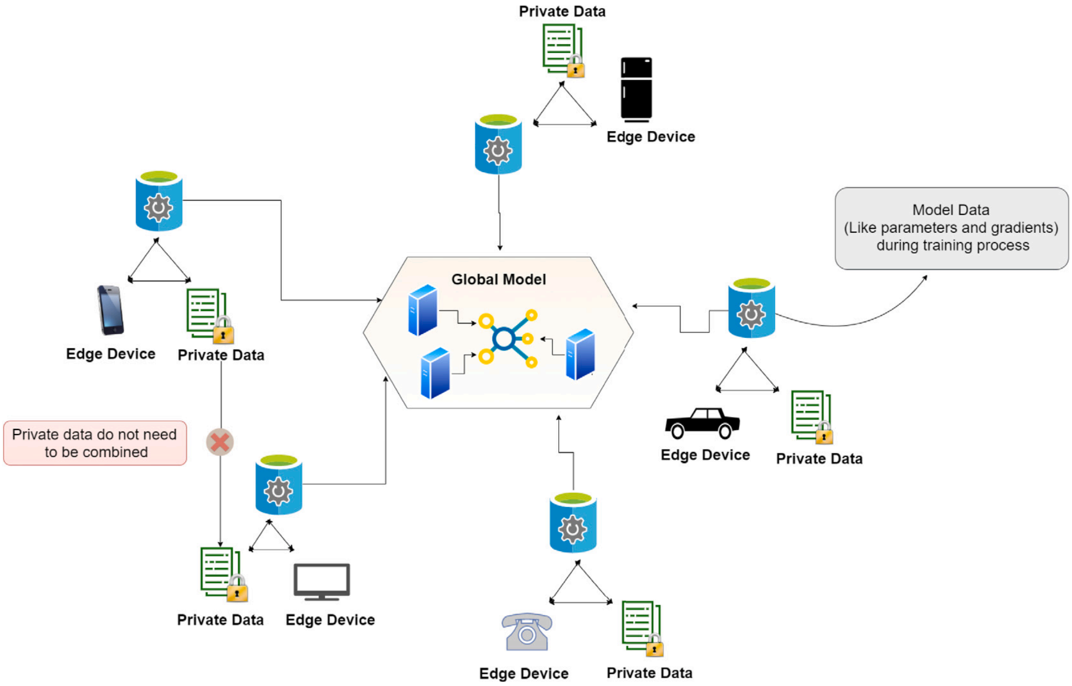


Fig. 5. A Federated Learning architecture modularized for privacy and security considerations (Lu, Liao, Lió and Hui, 2020).

privacy setup. One primary concern that arises with homomorphic encryption is on who holds the secret key of the encryption scheme (Gentry, 2009).

- Trusted Execution Environments (TEE), unlike the previous methods, aims to transition a module of the FL process into a secured environment to run the needed computations. The main challenge under this setup is that current TTEs suffer from limited resources, where for example only access to CPU hardware is available, and are unable to secure against all types of possible attacks to the environment (Subramanyan, Sinha, Lebedev, Devadas, & Seshia, 2017).

FL may combine some or all of the aforementioned secure computations to enhance the security and privacy of client data and the FL collaborative learning process. Fig. 5 presents a FL architecture modularized for privacy considerations. We discuss privacy in FL in more detail under Section 4.

3.2.4. *Heterogeneous devices*

In a FL setting where learning is decentralized, worker performance is not necessarily homogeneous. Device heterogeneity leads to dynamic computation capacity and communication conditions across clients. To solve this problem, new FL frameworks, such as heterogeneous FL and asynchronous FL (Sprague et al., 2018), address the issue by looking to provide heterogeneous clients with unique computational and communication capabilities. The authors of Chai et al. (2019) discuss how client training time is impacted by the available computational resources and the nature of data heterogeneity. Furthermore, the authors of article (Sarikaya & Ercetin, 2020) use a FL algorithm to reduce the training time of deep neural networks with high numbers of layers that are trained on large local datasets.

3.3. *Federated Learning architectures*

Federated Learning (FL) can be categorized into three architectures: Vertical Federated Learning, Horizontal Federated Learning, and Federated Transfer Learning. In this section, we describe each architecture and under which settings they are best utilized.

3.3.1. *Vertical Federated Learning*

Vertical Federated Learning, also known as *feature-based FL*, is typically used in settings where two (or more) client datasets share a similar sample ID space but have different input feature spaces. This type of learning allows clients to aggregates the feature information they have in combination for a particular sample ID, by using a third party that would secure that no information is shared about the unique sample ID during the process of feature sharing. Thus, vertical FL computes the cost function and gradients of a Machine Learning (ML) model while preserving the privacy of the unique sample ID data while collaboratively sharing the sample's unique features collected under different clients (Yang et al., 2019a).

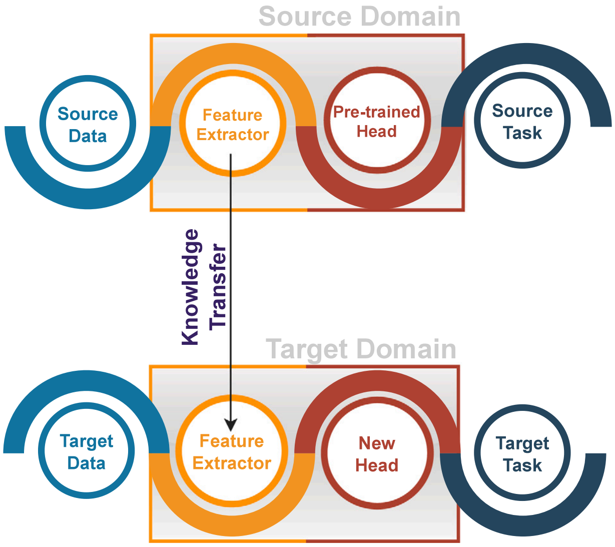


Fig. 6. A Transfer Learning architecture for head retraining.

As an example, assume we have separate companies (that is, the FL clients) located in the same city: a bank and an e-commerce company. The users' datasets for both companies are highly intersecting, meaning there is a large portion of shared users across the company datasets. The features collected for the two company datasets are different, representing different feature spaces, where for example the bank records the expenditure, credit rating, and the user's revenue while the e-commerce company records the user's browsing history and the purchase history of the user. With such a setting, it would be very useful to leverage vertical FL, where various companies with different datasets can build a collaborative ML model to achieve a richer learning of user behaviors (Song, Tong, & Wei, 2019). Vertical FL systems will treat the uniqueness and status of each participating user similarly and help develop a system that clients could share with ensuring the security and privacy of the clients' data.

Vertical FL assumes honesty but engages curious clients. It assumes that when two clients are involved, one is compromised of an adversary and the latter can only get data from the corrupted adversary client without being infected by it. The primary focus of vertical FL is privacy as its primary benefit lies in encryption.

3.3.2. *Horizontal Federated Learning*

Horizontal Federated Learning, also referred to as sample-based FL, is type of learning architecture that is introduced in the event that client datasets have a different sample ID space but share a similar feature space. For example, assume two local e-commerce businesses that have different users in their datasets from different regions but share a similar feature input space by collecting similar information about their users (Yang et al., 2019a). The different users in the e-commerce stores (that is, the FL clients) represent two different datasets or samples for a similar service being provided by the client, represented by a similar feature space.

The aforementioned scenario describes a setting that is suitable for a horizontal FL setup, where horizontal FL takes responsibility for honest partakers in sharing information of cross-users in different clients to provide a richer dataset for the ML model. Horizontal FL can also be used in a multi-tasking system, where multiple clients are allowed to learn separate tasks while sharing knowledge across their different samples and protecting the security of private information in the datasets. Horizontal FL ensures that no data leakage occurs during the sharing process across clients, and thus, data is protected and preserved. According to Aledhari et al. (2020a), horizontal FL focuses on security and its primary benefit is allowing for independence in learning across clients.

3.3.3. *Federated Transfer Learning*

Federated Transfer Learning (FTL) is applied to scenarios where two client datasets have a very small shared sample ID space with different feature spaces, somewhat similar to the vertical FL but over a smaller sample of sample ID intersection. Vertical FL is only applicable for an entire intersecting dataset sample space, so FTL provides an intermediate solution through Transfer Learning (TL) that allows to learn across the entire dataset even if only a small intersection exists across a similar sample space.

The standard TL approach of head retraining is shown in Fig. 6, where two different datasets exist in the process (1) a source dataset that trains an initial model that is composed of a feature extractor and a prediction head and (2) a target dataset that uses the previously trained (or 'pre-trained') feature extractor from the source dataset and trains a new prediction head. By leveraging the pre-trained feature extractor, the target dataset allegedly is able to improve its performance accuracy of the ML model, since usually the target dataset consists of fewer data samples relative to a rich source dataset. As such, FTL allows the transfer of knowledge across datasets of clients with small overlapping sample spaces (e.g. users) to build more accurate models.

Table 4 shows a comparison of the three category architectures of FL. Each category follows a different method for building an improved global model that allows for secure data and information sharing across clients. In summary, vertical FL merges the

Table 4
Federated learning architecture categorization.

Category	Definition	Use case	Security features
Vertical federated learning	Merge different features to achieve a richer feature space for ML models	Two datasets share a similar sample space with different feature spaces	Assumes honesty, Engages curious clients
Horizontal federated learning	Merge different samples to achieve a richer sample space for ML models	Two datasets have different feature spaces but share same sample space	Assumes honest partakers, Secure against curious service providers
Federated transfer learning	Build accurate target domain models by learning from rich source domains	Two datasets have different feature spaces but share small sample space	Security is infeasible

feature space across a similar sample space, horizontal FL in contrast shares different sample spaces for a shared feature space, and finally, FTL allows for sharing of knowledge in the case where only a small intersection of the sample space is shared with a different feature space.

3.4. *System challenges in Federated Learning*

We present three key challenges faced under a Federated Learning (FL) setting, specifically focusing on challenges arising in cross-device learning. The challenges we discuss are: reliability concerns in edge devices, risk of imbalanced data due to the nature of client selection across training rounds, and communication costs that dictate the broadcasting and receiving of model updates across clients and the central service provider.

3.4.1. *Reliability of edges devices*

Smartphone device battery life depends entirely on real-time communication and usage data. As a result, if a large number of connections are enabled, the battery of the device will drain faster than expected and be unable to handle all the running on-device activities. As a result, training models on edge devices causes a massive intake on the device's battery life.

The authors of Xu, Li, and Zou (2019a) show that applying a two-layered strategy on battery power devices can train a model with less energy consumption and sufficient accuracy. Furthermore, the authors of Yan, Chen, Feng, and Qin (2020) aimed to reduce communication during the training process of Federated Learning (FL) in order to maximize energy efficiency. By using Wasserstein Generative Adversarial Networks (WGANs) in their work, the data buffer is replenished recursively until the completion of the entire training rounds. The authors of article (Li, Xiong, Guo, Wang, & Xu, 2019) proposed a hierarchical online pace control framework, called SmartPC, to achieve an energy-efficient model training process by balancing the training time with the performance accuracy. Additionally, the authors of Tran, Kaddoum, Elgala, Abou-Rjeily, and Kaushal (2020) suggested implementing FL in wireless networks to notably reduce the lifetime of energy-constrained mobile devices. Therefore, they discussed a new approach at the physical layer based on the application of lightwave power transfer in a FL-based wireless networks as well as a resource allocation scheme to manage the network's power efficiency. In Yang, Chen, Saad, Hong and Shikh-Bahaei (2019), the authors analyzed the issue of energy consumption for task computation and transmission. They provided a solution through an optimization problem that utilizes a SVM-based FL algorithm.

3.4.2. *Imbalanced data*

Since the local data of clients are distributed across different participating edge devices, the problem of imbalanced data is very likely to occur leading to a significant impact on the performance of the global model by slowing down the convergence rate, increasing the model's bias, and decreasing the model's accuracy.

The authors of van Berlo, Saeed, and Ozcelebi (2020) considered training Deep Learning network models with limited labeled data in decentralized systems. They utilize pre-trained networks to enable superior feature extraction when dealing with fewer labeled data samples. Moreover, they propose a Deep Learning network that incorporates federated unsupervised representation learning as an improvement to traditional decentralized systems. In article (Duan et al., 2019), the authors proposed a solution to solve the issue of imbalanced distributed training data, which causes accuracy degradation in FL settings, by introducing a framework called Astraea. Astraea eliminates global imbalance by runtime data augmentation. In order to calculate average local imbalances, their method creates a mediator to reschedule the training of clients based on the Kullback–Leibler divergence (KLD) across the different client data distributions.

3.4.3. *Communication costs*

During each training round of FL, a set of interactions will occur between the participating clients and the central server to transmit local updates and broadcast copies of global models. This may cause network congestion and increase your overall communication costs.

The authors (Kharitonov, 2019) proposed FOLtR-ES, a learning algorithm that combines four principle requirements including preserving user privacy, minimizing communication and computation costs, using noisy bandit feedback to improve learning, and learning with non-continuous ranking quality measures. Even though many of the privacy concerns have been resolved by using an FL framework, communication costs remain a big constraint. The authors of Yao, Huang, and Sun (2018) proved that a two-stream model with Maximum Mean Discrepancy constraints reduces the communication costs of FL training rounds by 20%.

Resisting noise in wireless communication is another communication challenge faced in FL settings. In order to solve this problem, the authors of Ang et al. (2020) proposed a robust design architecture for FL to eliminate communicated noise. The authors believed that the noise issues can be viewed as a parallel optimization problem. In another attempt to tackle the reduction of client–server communication costs, the authors of Chen, Sun and Jin (2019) proposed an asynchronous model update strategy and a weighted aggregation method. The results showed that the proposed architecture outperformed baseline algorithms. In order to combat the issues seen in the standard Federated Averaging (FedAvg) algorithm, including both communication and performance issues, the authors in Yao, Huang, Wu, Zhang, and Sun (2019) proposed two solutions. The first solution was FedMMD that uses a two-steam model with Maximum Mean Discrepancy instead of using a single model like FedAvg to be trained on devices. The second is solution was FL with FedFusion that aggregates the features from local and global models resulting in higher accuracy and lower communication cost. The authors of Luping, Wei, and Bo (2019) provide a solution for the communication overhead problem in FL by proposing an algorithm called Communication Mitigated Federated Learning (CMFL), which is used to eliminate irrelevant client-side updates that are trained over client-specific, biased data.

At a high level, the authors of Li, Sahu, Talwalkar and Smith (2020) described some of the main challenges in FL, which include the following: expensive communication, systems heterogeneity, statistical heterogeneity, and privacy concerns. Many methods have been developed in order to conquer the issues of communication specifically. However, these solutions remain unable to fully handle federated networks. Additionally, the authors went on to describe future research direction in FL, which focus on extreme communication schemes, communication reduction and the Pareto frontier, novel models of asynchrony, heterogeneity diagnostics, granular privacy constraints, training methods that extend beyond supervised learning, productionizing FL, and appropriate FL benchmarking.

4. Privacy considerations in Federated Learning

4.1. *Privacy concerns in Federated Learning*

Trustworthy computing and networking systems are a long term objective for service providers (Jararweh, Otoum, & Al Ridhawi, 2020). The authors of Li, Sharma and Mohanty (2020) discussed the most important challenges related to preservation of client data privacy in FL settings including the attack mechanisms and associated solutions. The authors in paper (Yang et al., 2019a) presented that Artificial Intelligence, in general, still faces two major challenges. One, that exists in most industries, is that the data exists in the form of isolated islands. The other is in the reinforcement of data privacy and security. Thus, they wrote some of the possible solutions to those challenges, which including leveraging FL settings to counterpart the need to maintain data sources on independent islands while resolving any privacy and security concerns in the process.

The authors of Bhagoji, Chakraborty, Mittal, and Calo (2019) showed how FL model training is vulnerable to data poising, often referred to as model poisoning as well. Unlike traditional model settings, model poising is a small malicious aim to destroy the global model by having it misclassify specific inputs that lead to negative effects of other participating clients. One of the Federated Learning (FL) primary challenges is the malicious participation of clients who might inject the model with false input with the purpose of corrupting the global model. The authors of Chen et al. (2020a) designed a training-integrity protocol for Trusted Execution Environment to defect malicious attacks early. The authors of Mowla, Tran, Doh, and Chae (2020) provided a FL-based, on-device jamming attack detection security architecture for flying ad-hoc networks. The method is able to better identify client groups when the global model is being updated. In recent years, credit card fraud has caused a huge loss to both banks and consumers. Therefore, the authors in paper (Yang, Zhang, Ye, Li and Xu, 2019) presented a framework called Federated Learning for Fraud Detection, which trains a fraud detection model using behavioral features using FL.

4.2. *Privacy-preserving Federated Learning*

Although FL preserves user privacy by decentralizing data from the cloud server to live on end devices and solving data governance and ownership issues, it does not fully guarantee security. For example, attackers may steal personal data directly from edge devices, attack the communication process, or compromise the global model training process, where updates are attacked to not allow for secure aggregation.

Client-side Machine Learning (ML), privacy and security, interactive ML, and distributed bagging are four main building blocks of the proposed architecture for FL that the authors developed in Malle, Giuliani, Kieseberg, and Holzinger (2017) propose. They presented the concept of a local sphere, which represents a portion of the globally available information within a system. A possible workflow is also presented in their work (Malle et al., 2017), which describes the server as a global sphere, the client as a local sphere, and the users that are represented by data as the “world”. In Xu, Li, Liu, Yang, and Lin (2020), VerifyNet is used by authors to solve two issues of Deep Learning networks. The first issue was preserving the privacy of user information during the training process, and the second was verifying the accuracy of the model outcomes (or predictions) that are broadcasted by the server. The authors (Xu et al., 2020) proposed FedLDA that applies FL in Latent Dirichlet Allocation frameworks to reduce data collection risks. By training the model on three open datasets, their work demonstrated effectiveness in data privacy, model accuracy, and reduced communication cost.

In Lu, Liao et al. (2020), the authors discussed a Privacy-Preserving Asynchronous Federated Learning Mechanism for edge network computing (PAFLM), which achieved more efficient FL training without sharing private user data. The authors of Triastcyn and Faltings (2019) mentioned that Bayesian differential privacy provided sharper privacy loss bounds. Also, they recommended multiple improvements for more efficient privacy budgeting at different stages in the FL training process. The authors of Qian et al. (2019) discussed a Privacy-aware Service Placement (PSP) scheme to address the issue of service placement with privacy-awareness in edge cloud systems or cloudlets. Furthermore, the authors in Hu, Guo, Li, Pei, and Gong (2020) discussed a privacy-preserving approach for learning effective personalized models on distributed data through ensuring the differential privacy of user data. Also, they discussed tentative results on realistic mobile sensing data, which was shown to be robust to user heterogeneity and showed a good trade-off between accuracy and privacy.

In Huang et al. (2020) and Truex, Liu, Chow, Gursoy, and Wei (2020), the authors presented novel frameworks for Differential Privacy. In Huang et al. (2020), for example, the authors proposed a novel Differentially Private Federated Learning (DP-FL) framework for imbalanced data scenarios. (DP-FL) Furthermore, they presented how the DP-FL framework works in the cloud. The authors of Zhang, Wang, Zhao, and Chen (2019) presented an efficiently private FL scheme in Mobile Edge Computing (MEC), called FedMEC, to solve several issues in MEC systems. A Deep Learning network was trained on a dataset that can be exploited to partially reconstruct the training samples of the original dataset. In article (Zhao et al., 2019), the authors showed that the attackers can inject poisonous attacks to the FL model training round by uploading malicious model updates. Therefore, they suggested a novel Poisoning Defense Generative Adversarial Network (PDGAN) to defend against poisonous attack within FL settings.

The authors of Aïvodji, Gambs, and Martin (2019) present some security and privacy aspects by combining FL with secure data aggregation to provide high security to the connected client devices, which lead to an added level of convenience to the clients instead of making their information vulnerable to malware. The authors of Hao, Li, Xu, Liu and Yang (2019) integrated homographic encryption with different privacy techniques by using gradient descent as an efficient and privacy-preserving federated Deep Learning protocol to leverage the privacy of clients' data. Additionally, the authors of Wang et al. (2019) proposed a solution for privacy leakage of FL models. Unlike traditional FL that operates on the client-side, they use Generative Adversarial Networks with a multi-task discriminator to enable the retrieval of private client information and the invisible updates on the server-side. In Xu, Baracaldo, Zhou, Anwar and Ludwig (2019), the authors proposed HybridAlpha, an approach for privacy-preserving FL that employs an SMC protocol that supports functional encryption. The authors of Feng, Rong, Sun, Guo, and Li (2020) proposed PMF, a Privacy-preserving Mobility prediction framework via FL, to solve security concerns without significantly sacrificing the model's accuracy.

Researchers in Liu, Li, Xu, Lu and He (2020) have noted that if security is high in differential privacy settings, then accuracy is supposed to be sacrificed. Therefore, they propose an Adaptive Privacy-preserving Federated Learning framework (APFL) that maintains the privacy and accuracy of the trained models in FL settings. The researchers note that many current FL applications use either differential specificity or Secure Multiparty Computation (SMC) computations to ensure a secure FL framework. Hence, these techniques lack either accuracy or security. In Truex et al. (2019), the authors displayed an alternative approach that utilizes both differential privacy and SMC to balance between privacy and accuracy. Moreover, the authors of Fang, Guo, Wang, and Ju (2020) implemented an encryption protocol to provide enhanced privacy preservation, in addition to preserving the usefulness of the desired model. Moreover, they proposed strategies to improve training efficiency. The authors in Chen et al. (2020b) proposed a new privacy-preserving FL framework to guarantee the integrity of Deep Learning training processes, where they utilized Trusted Execution Environments (TEE) for a secure computation environment in the FL setting. The authors of Bagheri, Rezapoor, and Lee (2020) proposed an encrypted federated framework that ensures the privacy of client data, while generating comprehensive models that leverage a variety of multiple enterprise operations. In Ma, Zhang, Chen, and Shen (2018), the authors identified some privacy-preserving issues in the FL data training process, so they propose a new framework for privacy-preserving multi-party Deep Learning in FL.

In article (Sariyildiz, Cinbis, & Ayday, 2020), the authors proposed a novel classification model that is resilient against attacks by design. Moreover, they showed how to use high dimensional keys to develop robustness against attacks without increasing the model complexity. The authors in Liu et al. (2020) studied a privacy-preserving federated k-means scheme (PFK-means) for proactive caching within the next-generation cellular networks. PFK-means is founded on a privacy-preserving methods that relies on the combination of FL and secret sharing. In Nishio and Yonetani (2019), the authors discussed a decentralized learning framework that enables privacy-preserving training of models to work with heterogeneous clients in a practical cellular network. They used a FL protocol known as FedCS and found that it is able to complete an FL training round in a significantly shorter time compared to the original FL protocol. The authors in Dong, Chen, Shen, and Wang (2020) analyzed the privacy leakages of TernGrad. They also explore EaSTFLy, a solution to solve the privacy leakage problem. Hence, they presented how EaSTFLy works and analyzed its performance when applied on TernGrad.

The authors of Li and Han (2019) displayed an end-to-end encrypted Deep Learning network architecture, which is being used for compressing and encrypting gradient updates. The goal of their method was to ensure that gradient updates that are encrypted are not disrupted during the process of communication and not revealed to the server. The authors in Zhang, Fu, Wang, Zhou, and Chen (0000) build out a privacy-preserving and verifiable FL scheme, where they introduce a bilinear aggregate signature technology into the FL framework. Moreover, the authors in Taïk and Cherkaoui (0000) analyzed the usage of Edge Computing and FL. They showed that their proposed method is the first to utilize FL household power load forecasting with promising results. The authors in paper (Liu, Li, Xiao and Jin, 2019) proposed FLoc, a WiFi indoor location fingerprint localized system based on Federated Learning. One of the key features of their system was that it focused on reducing privacy breaches when updating the fingerprint-based localized model.

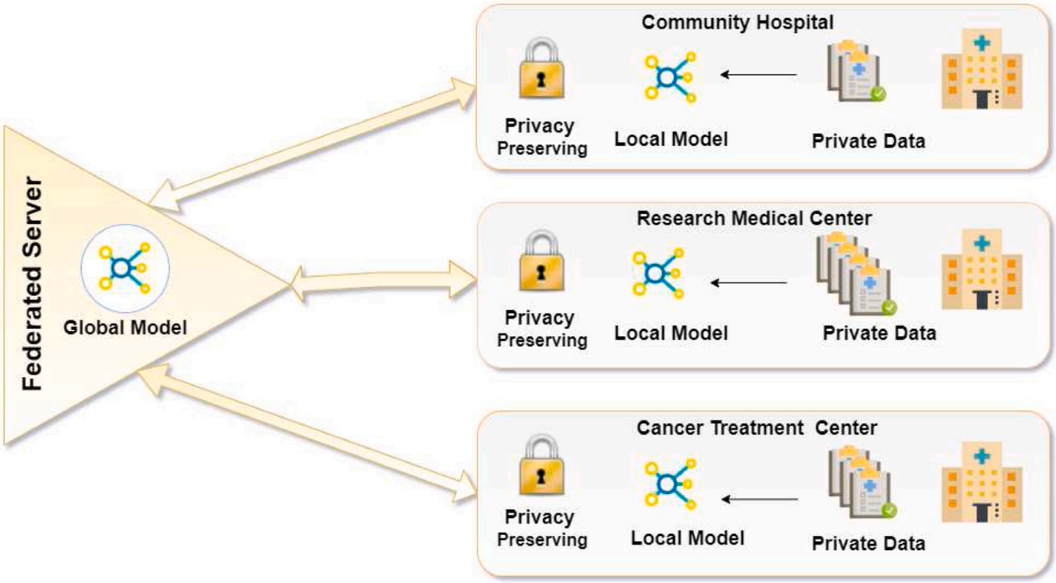


Fig. 7. Privacy-preserving Federated Learning within a personal healthcare framework (NVIDIA Blog, 2020).

Despite FL's numerous benefits, it has attack vectors that adversaries can utilize which present major challenges and drawbacks to the FL framework. The authors in Nuding and Mayer (2020) conducted a research study about the manipulations and challenges that may occur in the training process of FL settings. In Ma et al. (2020), the authors discussed the privacy and security issues that arise in FL and several challenges to preserving privacy and security of participating clients as well as potential solutions to the raised issues. The authors of Wang, Tong and Shi (2020) designed a technique to reduce the communication cost between the data collector and the clients during model training. The authors of NVIDIA Blog (2020) provided a comprehensive architecture for privacy-preservation in FL for personal healthcare applications, that is presented in Fig. 7.

5. Applications & trending use cases

Federated Learning (FL) brings collaborative Machine Learning (ML) to industries to gain more benefits from an extensive variety of distributed datasets, accelerate various industrial processes, and support privacy-sensitive applications. One of the most utilized FL applications by the public is likely Gboard (Hard et al., 2018), Android's Google keyboard, for the purpose of predictive texts. Another example of FL applications is in making it possible to leverage data islands across healthcare institutions to improve secure cross-data analysis. FL is usually used by researchers to train models on client edge devices without relocating the data as it can be too sensitive and private or due to the difficulty of collecting the data from different clients. Furthermore, FL is applied to a wide scene of website applications such as entertainment, dating, and e-commerce to provide personalized recommendations depending on the interests of the customers by exploiting similarities in customers' interaction patterns (Yang et al., 2019b). Moreover, with autonomous vehicles, each self-driving car generates a huge amount of data, which may potentially lead to communication delays in the response time and needed action to be taken by the vehicle during data communication over the network. In cases such as the aforementioned, FL is a recommended solution, where each client trains a model locally and then transfers a compressed version of new information, or parameter updates to the global model at the server end (Lu, Huang, Dai, Maharjan and Zhang, 2020).

5.1. Federated Learning in technologies

5.1.1. Federated Learning in artificial intelligence

In article (Zhou, Yang, Pu, & Yu, 2020), the authors discussed how to coordinate across edge devices and the cloud to optimize the system-wide cost efficiency of FL. By using the Lyapunov optimization theory, they designed and analyzed a cost-efficient optimization framework (CEFL). The authors proposed a FL framework to address data shortage and security issues in industrial applications of AI. In the study (Han, Yu, & Gu, 2019), the authors discussed the workability of their proposed FL framework and the expectations that this framework brings in terms of safety, reliability, and efficiency. The authors of Hao et al. (2019) proposed an efficient Privacy-Enhanced Federated Learning (PEFL) scheme for industrial AI and compared it with existing solutions in literature.

Additionally, the authors in Wang et al. (2019) designed the "In-Edge AI" framework in order to intelligently utilize the collaboration among devices and edge nodes to exchange the learning parameters for better training and inference of the models. Also, they discussed how to integrate Deep Reinforcement Learning techniques and FL frameworks with mobile edge systems.

In Verma, White, and de Mel (2019), the authors described a web-service based implementation of FL systems. They focused on the problems that companies face while using distributed data and discussed some potential solutions. The authors of Sozinov, Vlassov, and Girdzijauskas (2018) evaluated FL to train a human activity recognition classifier and compared its performance to centralized ML.

Moreover, the authors in Shan, Cui, and Gao (2020) proposed an intelligent resource allocation model "DRL+FL" that can solve problems related to uploading large amounts of training data via wireless channels, non-IID and imbalance of training data when training DRL agents, restrictions on communication conditions, and data privacy. In Zhao, Chen, Wu, Teng and Yu (2019), the authors proposed a multi-task Deep Learning network in FL (MT-DNN-FL) in order to run network anomaly detection. The authors of Liu, Wang and Liu (2019) presented a learning architecture for navigation of cloud robotic systems: Lifelong Federated Reinforcement Learning (LFRL). This algorithm is helps update a shared model deployed on the cloud.

5.1.2. *Federated Learning in Internet of Things*

There are many issues that arise within the deployment process of FL in resource-constrained Internet of Things (IoT) environments. In an attempt to solve these issues, the authors of Feraudo et al. (2020) have proposed CoLearn, an architecture that is built on top of the open source Manufacturer Usage Description (MUD) implementation (osMUD) and the FL framework PySyft. In Sun et al. (2020), the authors created the General Gradient Sparification (GGS) framework designed for FL in Edge Computing environments. Moreover, they conducted experiments on LeNet-5, CifarNet, DenseNet-121, and AlexNet with adaptive optimizers. In article (Wu, He, & Chen, 2020), the authors proposed a framework called PerFit for personalized FL that eliminates devices' heterogeneous nature, statistical data heterogeneity, and model heterogeneity in IoT applications. Furthermore, they provided a case study of IoT-based human activity recognition to demonstrate the effectiveness of personalized FL for intelligent IoT applications.

Most of the existing research within the IoT space focuses on deployment of additional sensors for data collection. In Imteaj and Amini (2019), the authors propose a framework to create sensors that are built into smartphones for decision making in different areas. The idea behind this framework is that each device serves as a distributed decision making structure. The authors of Kwon, Jeon, Park, Kim, and Cho (2020) proposed an approach to incorporate FL-based distributed Deep Learning into Internet of Underwater Things (IoUT) networks by using a MADDOG-based algorithm with small sized iterations. Compared to reinforcement learning based on JCARA methods, the MADDOG-based algorithm achieved better performance. The authors of Zhou, Li, Chen, and Ding (2018) proposed a real-time data processing architecture for multi-robot environments based on differential FL. The architecture proposed acquires purposeful information while maintaining real-time data processing as well as data privacy.

FL faces issues with centralized optimization that relies on a central server, and this can lead to a single point of failure as well as scalability issues within the network. The authors in paper (Savazzi, Nicoli and Rampa, 2020) created a fully distributed FL algorithm that creates data functionalities inside the network to eliminate the reliance on a central server and avoid single-points of failure. In Samarakoon, Bennis, Saad, and Debbah (2020), the authors addressed the problem of joint power and resource allocation (JPRA) for low latency communication in vehicular networks. In order to solve this problem, they used a FL-based joint transmit power and resource allocation framework in order to achieve low-latency in vehicular communication. The authors of Hsu, Srivastava, Wu, and Chen (2020) studied the industrial IoT space to create data processing and analysis techniques in computing environments by using Deep Learning technology to predict component status and life. This is conducted by using a RUL prediction-based approach. The authors in Posner, Tseng, Aloqaily, and Jararweh (2021) provided a comprehensive analysis of opportunities and solutions for FL in vehicular networks. Additionally, the authors of Du et al. (2020) discussed FL in IoT at a high level. Their work is divided into 6 main points of discussion which are basics and advancements in FL, technical challenges and solutions of FL in wireless IoT environments, advantages and technical issues in FL in vehicular IoT, and future research in FL. Future works state that by using FL, the training results within IoT applications can a richer user experience (Hsu et al., 2020).

5.1.3. *Federated Learning in blockchain*

Blockchain applications are expanding rapidly and covering many new domains (Tseng, Yao, Otoum, Aloqaily, & Jararweh, 2020). As a result, FL and blockchain applications are getting more and more attention from the research community. The authors of Doku, Rawat, and Liu (2019) proposed an approach to create a network sharing technique called Interest Groups. Nodes that share the same interests are categorized to the same Interest Group. Additionally, the data is stored centrally and, through using a proposed consensus mechanism called Proof of Common Interest, the data is being validated in order to ensure that it is relevant. In Kim, Park, Bennis, and Kim (2019), the authors proposed a blockchain-based FL (BlockFL) architecture, where the local learning model updates are traded in and authenticated. In addition, the authors analyzed the end-to-end latency model of BlockFL and determined the block generation rate by looking at the communication, consensus delays, and computation costs.

In Pokhrel and Choi (2020a), the authors proposed a blockchain-based FL (BFL) design for privacy-aware and vehicular communication networking with on-vehicle ML model updates being distributed. One of the key features of BFL is the enablement of on-vehicle ML without the constraint of centralized training data and without the need to use a consensus mechanism of the blockchain. The authors in Marulli, Bellini, and Marrone (2020) intended to provide a strong architectural base in order to solve the issues that are present in federated models. They suggest that the issues are resolved by using Cloud Computing technology that provides potential solution functionalities. Additionally, at a high level the paper touched base on the technical side of blockchain to improve the security of FL in general. The authors in Sharma, Park, and Cho (2020) proposed a distributed computing defence framework for sustainable society using blockchain and FL. The model utilized an algorithm to address the problems of limited training data to maintain high accuracy. The results showed that the proposed model is sufficient enough to train high accuracy models.

Using AI as a big data analysis tool has some challenges including centralized architecture, security measures, resource limitations, and insufficient training data. Incorporating blockchain will enable a decentralized architecture and integration of FL will improve scalability concerns. In Singh, Rathore, and Park (2019), the authors proposed the implementation of a Block IoT Intelligence architecture to eliminate the challenges in AI. 5G systems also face a number of security problems. Therefor, the authors in Nguyen, Pathirana, Ding, and Seneviratne (2019) created a comprehensive survey to describe the usefulness of integrating blockchain with 5G networks. Focusing on the advantages that blockchain has on different areas, which include spectrum management, data sharing, resource management, and FL. They presented methods of blockchain-based node recognition in order to rapidly improve the learning speed. The algorithm the authors proposed in Kim and Hong (2019) proved to be much faster and stabilized compared to the standard FL algorithm. In Majeed and Hong (2019), the authors proposed a blockchain network-based architecture called FlChain in order to improve on the security aspect of FL, using the concept of channels for learning many global models on FLchain. Qualitative analysis displayed that FLchain is a stronger framework than the traditional FL scheme.

Moreover, the authors of Qu et al. (2020) proposed a blockchain-enabled FL (FL-Block) scheme that enables local learning updates of end devices and exchanges with a blockchain-based global learning model, which is approved by blockchain miners. Additionally, the FL-Block allows autonomous ML with a central authority to look over the global model using the proof-of-work feature of the blockchain. FL, in return, allows participants to donate their local data without exposing it. In Martinez, Francis, and Hafid (2019), the authors implemented an EOS blockchain design to ensure data privacy. The authors of Awan, Li, Luo, and Liu (2019) presented the implementation of a blockchain-based privacy-preserving FL framework to show the possibility of updating the global model parameters securely in a FL setting by using the blockchain decentralized trust properties.

In paper (Lu, Huang, Dai, Maharjan, & Zhang, 2019), the authors designed a blockchain authorized secure data sharing architecture for distributed multiple parties. Then, they framed the data sharing problem into a ML problem by incorporating privacy-preserved FL because the seepage of private data can lead to serious problems beyond financial damage for the providers. The authors of Toyoda and Zhang (2019) proposed a generic full-fledged protocol design for FL on a public blockchain to realize desired objectives under the assumption that participants act rationally. Furthermore, they theoretically clarified incentive compatibility based on contest theory of an auction-based game theory in economics. The authors in the study (Liu et al., 2019) adopted blockchain as the ML environment, where different actors collaborate on the training mission. Throughout the training process, an encryption algorithm is employed to shield the privacy of information and the trained global model.

5.1.4. Federated Learning in Natural Language Processing

The study (Zhu, Wang, Hong, Xia, & Xiao, 2019) presented the aimNet network to solve vision-and-language grounding problems. The proposed network was applied for image captioning and Visual Question-Answering (VQA) tasks. Both tasks showed improvement in performance. The authors also showcased the function of aimNet under the three FL settings: horizontal FL, vertical FL, and Federated Transfer Learning (FTL). Moreover, in Liu, Wu, Ge, Fan and Zou (2020), the authors applied FL to a deep convolutional network to perform variable-length text recognition over a large corpus. Furthermore, they conducted a comparison between two FL frameworks, TensorFlow Federated and PySyft, in order to see which one achieved higher accuracy. The authors of Lin et al. (2021) proposed FedNLP, an NLP platform that effectively employs FL techniques for automated natural language tasks such as text classification, sequence tagging, and question-answering amongst others, to train Natural Language Processing (NLP) models in a distributed computing environment. The work (Hilmkil et al., 2021) examined FL settings for tuning transformer language models. It also evaluated three variants of the BERT transformer, BERT, ALBERT, and DistilBERT, for different NLP tasks such as text classification and sentiment analysis tasks. Additionally, the authors of Lin, Kong, Stich, and Jaggi (2020) presented a flexible aggregation scheme of ensemble distillation for FL models. Their proposed scheme improved privacy, reduced cost, and allowed resilient integration over heterogeneous client models. They conducted comprehensive experiments on different tasks including NLP.

5.1.5. Federated Learning in autonomous vehicles

In Samarakoon, Bennis, Saad, and Debbah (2019), the authors studied the issue of JPRA for ultra-reliable low-latency communication (URLLC) in vehicular networks. They discussed the issue of network-wide power consumption of vehicular users that is minimized subject to high reliability in terms of probabilistic queuing delays through using many theories. After the authors in Lu, Huang, Dai et al. (2020) detected that data leakage in VCPS can lead to physical consequences, for example, endangering passenger safety, privacy, and causing severe property loss for data providers. They proposed a solution using FL to secure data privacy. The authors of Pokhrel and Choi (2020b) proposed a new communication efficient and privacy-preserving FL framework for improving the execution of the Internet of Vehicles (IoV). The authors in Lu, Huang, Zhang, Maharjan and Zhang (2020) saw that allowing IoV environments to share data across vehicles for collaborative analysis can greatly improve the driving experience and service quality and resolve issue of intermittent and unreliable communications in IoV. They proposed a new architecture based on FL to address privacy concerns in IoV environments.

The authors of Saputra et al. (2019) presented novel approaches utilizing state-of-the-art ML strategies pointing at foreseeing energy requests for electric vehicle (EV) systems. The authors in Pokhrel and Choi (2020a) enhanced an autonomous BFL design for privacy-aware and efficient vehicular communication networking. BFL has effectively enabled on-vehicle ML without any centralized coordination by availing the consensus mechanism of blockchain. The authors in Samarakoon, Bennis, Saad, and Debbah (2018) discussed ultra-reliable low-latency communication (URLLC) in vehicular networks. Also, they discussed the problem of joint power control and resource allocation for V2V communications. Moreover, the authors of Brik, Ksentini, and Bouaziz (2020) proposed Unmanned Aerial Vehicles (UAVs) that can act as mobile base positions to improve the capacity, coverage, and energy efficiency of wireless networks. The authors in Zeng et al. (2020) created a framework in order to use distributed FL algorithms with a UAV swarm. Compared to the baseline design, the simulation results validated the effectiveness of the FL framework through a convergence analysis.

Table 5
A summary of literature resources for technology application of Federated Learning.

Technology applications	Literature resources
Federated Learning in Artificial Intelligence	Arrieta et al. (2020), Han et al. (2019), Hao, Li, Luo et al. (2019), Liu, Wang et al. (2019), Shan et al. (2020), Sozinov et al. (2018), Verma et al. (2019), Wang, Han et al. (2019) and Zhou et al. (2020)
Federated Learning in IoT	Du et al. (2020), Feraudo et al. (2020), Hsu et al. (2020), Imteaj and Amini (2019), Kwon et al. (2020), Samarakoon et al. (2020), Savazzi, Nicoli and Rampa (2020), Sun et al. (2020), Wu et al. (2020) and Zhou et al. (2018)
Federated Learning and Blockchain	Awan et al. (2019), Doku et al. (2019), Kim et al. (2019), Liu, Hu et al. (2019), Lu et al. (2019), Majeed and Hong (2019), Martinez et al. (2019), Marulli et al. (2020), Nguyen et al. (2019), Pokhrel and Choi (2020a), Qu et al. (2020), Sharma et al. (2020), Singh et al. (2019) and Toyoda and Zhang (2019)
Federated Learning in NLP	Hilmkil et al. (2021), Lin et al. (2021, 2020), Liu, Wu et al. (2020) and Zhu et al. (2019)
Federated Learning in Autonomous Vehicles	Brik et al. (2020), Lu, Huang, Dai et al. (2020), Lu, Huang, Zhang et al. (2020), Pokhrel and Choi (2020a), Pokhrel and Choi (2020b), Samarakoon et al. (2018, 2019), Saputra et al. (2019) and Zeng et al. (2020)
Federated Learning in Resource Allocation	Chen, Chuang and Wu (2020), Conway-Jones, Tuor, Wang, and Leung (2019), Liu, Chen, Chen and Zhang (2020), Liu, Zhang, Song and Letaief (2019), Pandey et al. (2020), Sattler, Wiedemann, Müller, and Samek (2019), Savazzi, Nicoli, Rampa and Kianoush (2020), Shi, Zhou, and Niu (2019), Sun et al. (2020), Sun, Zhou, and Gündüz (2019), Yang, Liu, Quek and Poor (2019), Yu et al. (2020) and Zou et al. (2019)

5.1.6. Federated Learning in resource allocation

The authors in Sun et al. (2020) built a General Gradient Sparsification (GGS) framework that aims to solve the problem of sparse gradient update processes using two important procedures: gradient correction and batch normalization updates with local gradients. In FL, there is often a central server in the cloud or at edge, where each cloud component has its advantages and disadvantages. Edge servers, for example, provide more efficiency in communication with clients. The authors in Liu, Zhang et al. (2019) discussed combining both edge and cloud servers' advantageous features and proposed a client-edge-cloud FL system using the HierFAVG algorithm. In addition, the authors of Sun et al. (2019) proposed an online energy-aware dynamic worker scheduling policy. Experiments using MNIST dataset showed that for non-IID data, doubling data storage can improve the accuracy under an astringent energy budget. In Shi et al. (2019), the authors created a joint bandwidth allocation scheduling problem in order to understand the convergence performance of FL. Overall, FL was shown to ensure privacy-preserving collaborative learning. However, a communicational overhead was witnessed during the training process. In order to solve this issue, the authors in Sattler et al. (2019) proposed Sparse Ternary Compression (STC), a framework designed to target FL environments. STC was shown to outperform Federated Averaging (FedAvg). The authors of Chen, Chuang et al. (2020) created the federated extreme learning machine system (Fed-ELMS), which showed many added benefits to traditional FL including higher training speed, better performances, and efficient computations. These benefits are of great value to participating client edge devices. The authors in Xu, Li, and Zou (2019b) presented work to assess battery resources in edge devices. They analyzed the possibility of enabling FL on battery powered devices. They proposed a two-layered strategy. The first layer improved the initialization of FL while the second layer explored local energy saving potential.

Moreover, the authors in Savazzi, Nicoli, Rampa and Kianoush (2020) proposed a distributed FL framework that performs a decentralized fusion of local model parameters by connecting devices and local (ie, in-network) data through consensus-based methods. A benefit of this framework was the blueprint that the framework is able to create for future 5G wireless networks. A FL contributor receives a pay for participating in the FL model training rounds. The authors of Yu et al. (2020) developed a FL Incentivizer (FLI), which divides a pre-defined pay amount across data owners to help reduce the waiting time across participating FL contributors. In Liu, Chen et al. (2020), the authors proposed momentum FL (MFL) that makes use of momentum gradient descent in the local update step of FL settings. Their approach helped resolve the issues that arise in distributed ML frameworks. In Conway-Jones et al. (2019), the authors proposed a resource-constrained network environment that is decentralized and where the availability of clients is irregular. The authors of Zou, Feng, Xu et al. (2019) proposed a two-layer dynamic game model, which incorporates a lower-level evolutionary game of the model owners and an upper-level evolutionary game of mobile device groups. The results are verified through numerical evaluations.

One of the primary resource constraints with FL settings is how end users can collaborate to build a global model with communication efficiency over the wireless network. In Yang, Liu et al. (2019), an analytical model is developed to characterize the performance of FL in wireless networks. Using the developed analysis, the effectiveness of three different scheduling policies were studied, including random scheduling (RS), round robin (RR), and proportional fair (PF), and compared in terms of their FL convergence rate. The authors saw that a key challenge in FL was how users participate to build a high-quality global model while maintaining communication efficiency. Consequently, the authors in Pandey et al. (2020) decided to tackle this issue by formulating a utility maximization problem and proposed a novel crowd-sourcing framework to leverage the benefits of FL settings. See Table 5 for the summary of literature resources for technology application of Federated Learning.

Table 6
A summary of literature resources for market use cases of Federated Learning.

Market use cases	Literature resources
Federated Learning in Data Science	Duan et al. (2019), Muniswamaiah et al. (2019), Xu et al. (2019b), Zhao, Chen, Wu et al. (2019), Zhou et al. (2018) and Zou et al. (2019)
Federated Learning in Healthcare	Blanquer et al. (2020), Brisimi et al. (2018), Chen, Li, Xu, Zhang and Luo (2020), Chen, Qin et al. (2020), Deist et al. (2020), Elayan et al. (2021a, 2021c), Goecks et al. (2020), Huang et al. (2019), Kuai and Zhong (2020), Lin (2020), Lundervold and Lundervold (2019), Silva et al. (2019) and Strangman et al. (2019)
Federated Learning in Industry	Franco, Van, Dreiser, and Weiss (2021), Gengler (2019), Guo, Wang, Vishwanath, Xu and Li (2020) and Zellinger et al. (2021)
Federated Learning in Education	Balta et al. (2021), Encheva and Tumin (2008), Guo, Zeng and Dong (2020) and Qiang, Lixin, Richard, et al. (2021)

5.2. *Federated Learning in market use cases*

5.2.1. *Federated Learning in data science*

Despite FL becoming more popular, there has been limited applications of FL within the Data Science industry. Edge Computing systems' training datasets are typically imbalanced, which affects the accuracy in FL-powered applications. To solve this problem, the authors in Duan et al. (2019) proposed a framework to create a self-balancing FL framework called Astraea, which removes the imbalances in the global data distribution. The authors in article (Muniswamaiah, Agerwala, & Tappert, 2019) implemented a federated query processing framework that extracts data and stores it in a common in-memory format. This is useful because the framework helps improve the efficiency of data extraction in data science project pipelines.

5.2.2. *Federated Learning in healthcare*

In Brisimi et al. (2018), the authors use FL to predict the hospitalization rate of patients with heart disease, while the authors of Silva et al. (2019) recommended to use the FL to study the relationship between diseases and the structure of the brain on large datasets collected across different clinic. Both works (Brisimi et al., 2018; Silva et al., 2019) aimed to avoid exchanging the data of patients for security and privacy purposes. The author of Huang et al. (2019) trained a community-based FL model to predict the mortality and ICU stay time by using the data of patients within a similar clinic instead of across multiple clinics. The developed model resulted in very accurate results with similar performance to centralized modeling, while securing the patients' private data. The authors of Blanquer et al. (2020) built a Rheumatic Heart Disease classifier using a federated cloud architecture to train the model without relocating patient data. Their method increased the security of patients' private information, since the data is encrypted in memory and disk instead of being found on public cloud infrastructures or research centers. Additionally, the authors of Kuai and Zhong (2020) proposed a FL algorithm to compute the performance of training tasks on local edge nodes to solve the security concern that appears during the process of data-brain model improvement and development. The information of lung cancer patients from different countries has been used for training a distributed logistic regression model to predict post-treatment two-year survival (Deist et al., 2020).

In the study (Lin, 2020), the authors use Artificial Intelligence in the radiology field to improve patient care from a hybrid academic-industry perspective. FL is used in order to train the model locally while keeping preserving the privacy of patient information. Also, the authors of Silva et al. (2019) presented a FL framework for safely accessing and meta-analyzing any biomedical data without without the need to share the data over a network. Furthermore, the authors in Strangman et al. (2019) recommended using FL to build a model that helps preserve the health of astronauts. The model assessed the medical condition of the astronauts under the same conditions that occurred on Earth. In Lundervold and Lundervold (2019), the authors proposed to replace Deep Learning algorithms with FL for enhanced security and to remove the dependency on relocating patient data. The authors in Goecks, Jalili, Heiser, and Gray (2020) also suggested using FL within a biomedicine to train a model using data from different clinical systems, in order to maintain the security of patient data. In Chen, Qin, Wang, Yu and Gao (2020), a FedHealth framework was found to successfully aggregate patient data through FL in order to learn personalized patient models.

Furthermore, the authors of Elayan, Aloqaily, and Guizani (2021a) proposed a framework that employed Deep FL for healthcare decentralized systems that rely on IoT devices in their structure. They integrated Deep Learning with FL to build robust models for a better remote health monitoring system. In the study (Elayan, Aloqaily, & Guizani, 2021c), the authors presented solutions for the sustainability of healthcare IoT-based data analysis systems. They proposed a framework that preserves privacy for intelligent healthcare IoT-based system users and supports the decentralized structure of the data at the same time. Moreover, they experimented with a decentralized skin disease detection system using FL and presented new FL algorithms to support the sustainability of FL-based systems by automating the training data acquisition process.

5.2.3. *Federated Learning in industry*

The author of Gengler (2019) utilized FL to solve the challenges caused by sensing data used to monitor dairy production in cattle farms. The purpose was to value dairy cattle from a genetic perspective, as well as to increase profitability of the farmers' businesses. In Guo, Wang et al. (2020), the authors studied the performance of heating, ventilation, and air conditioning using FL to preserve

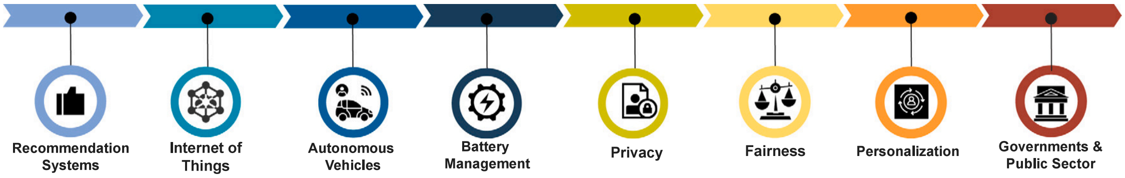


Fig. 8. Promising areas of future directions in Federated Learning research and applications.

the privacy of occupants' information. They designed BuildFL for FL of HVAC models and studied the impact of different factors on the accuracy of the model. The authors of Zellinger et al. (2021) investigated the gap between FL systems and the required industrial system settings. Therefore, they presented a module-based platform for designing and implementing transfer learning models for smart manufacturing systems that maintain the confidentiality of data. The authors of Franco et al. (2021) worked on automated industrial systems through a self-adaptive FL framework by training each participating factory model using a self-adaptive method to reduce communication overheads. The evaluation results showed a higher accuracy for the proposed architecture compared to the standard Federated Averaging (FedAvg) method.

5.2.4. *Federated Learning in education*

Guo, Zeng et al. (2020) present an FL-based education data analysis framework, FEderated Education Data ANalysis (FEEDAN), that breaks the borders of data islands when looking into pedagogical data analysis in efforts of pursuing the education 4.0 revolution. By using FL to eliminate institutions' concerns on student privacy, the authors (Guo, Zeng et al., 2020) open up possibilities for realizing large-scale educational analysis, which was not previously possible. Balta et al. (2021) present interesting work that studies accountability of the FL process and its implications on legislative and jurisdictional issues that generally arise in public sectors, where public educational institution may be involved. By studying the accountability framework of FL, more trust is brought forth to educational institutions in fostering the reality of a 4.0 educational revolution.

Reusable learning objects have important role in improving student learning experience. The reusable learning object allows the expert to share their views on a learning object by selecting one out of the several options in a given learning system. In Encheva and Tumin (2008), S. Encheva et al. develop a FL setting for learning model for reusable learning objects, which focus on responses collected from two experts. Furthermore, personalized learning and constraints of privacy preservation of participating parties (including students, teachers, and institutions) is another area that has been evaluated in the work of Qiang et al. (2021). See Table 6 for the summary of literature resources for market use cases of Federated Learning.

6. Future work directions

In this section, we discuss open future directions of Federated Learning (FL) with promising areas of growth, which are shown in Fig. 8 and include: Recommender systems, Internet of Things (IoT), autonomous vehicles, battery management, privacy, fairness, personalization, and governments and public sectors.

Federated Learning in recommender systems: There has been work that looked at incorporating FL into recommender systems, such as Dolui et al. (2019) and Jalalirad, Scavuzzo, Capota, and Sprague (2019). The aim in such studies and future investigations should focus on answering: (1) how to minimize user data sharing while maintaining accuracy, and (2) how optimization of FL frameworks may impact performance of standard recommender systems.

Federated Learning in IoT: Research in Internet of Things (IoT) is making use of FL to enhance data security for the participating edge devices (Rey, Sánchez, Celdrán, & Bovet, 2022). In such settings, FL trains ML models without needing to specific client data. With the advent of 5G, the future of FL in IoT will be susceptible to higher opportunities of cyberattacks. Hospitals, for example, have adopted FL to ensure data integrity and patient privacy while exchanging information about patient prognosis and diagnosis (Alawadi et al., 2021).

Federated Learning in autonomous vehicles: Autonomous vehicles make use of Machine Learning (ML) to detect changes in the road environment and learn how to respond to those changes. Beneath this, ML models are learned on the cloud provided by the automaker. The ML model is updated locally on the AI-powered self-driving car. The observed local data may include personal information that should not be uploaded to the cloud and that infringes the right to privacy (Posner et al., 2021). Thus, FL research is naturally emerging and growing under this area with a number of works studying FL settings for learning the various and complex ML models needed an autonomous driving setting.

Federated Learning in battery management: FL has been used in various edge devices to reduce the energy consumption of batteries for longer battery life and quality. ML modeling for edge devices tends to consume a lot of power, and thus, consumes the device's battery resources. The majority of battery consumption is mainly caused by continuous data transfers to and from the cloud. Thus, FL provides a resolution to development of ML models, where models are trained locally and local updates are then communicated to the cloud. The communication of the local updates is much lighter in communication cost, and this relieves many of resource allocation consumption in batteries. For future directions, more work on FL should focus on energy consumption of edge device battery life (Tang et al., 2021).

Federated Learning in privacy: FL aids privacy preservation by training local models to replace raw data exchange over the network. The updates of local models are sent to a centralized server to train a global model. The service provider has access rights to the server at regular intervals, where the global model lives and local updates are communicated. The future of privacy within FL will see the introduction of PYSYFT, which will be used to enable secure and private Deep Learning (Ziller et al., 2021). More work is to be explored on security measures of FL frameworks to avoid breaches of client data, local update communication and aggregation, and global model updates. Breaches to model training may results in serious repercussions such as biasing the model to favor the attackers personal agenda.

Federated Learning in fairness: Fairness in ML is typically defined using users' demographic information such as location, gender, and race (Oneto & Chiappa, 2020). In FL settings, demographic information of clients is not available, and more so, the one-to-one relationship that exists between an edge device and a user is not maintained. This makes measuring and correcting for unfairness in FL models infeasible in most cases (Kairouz et al., 2021). Many open questions remain to be studied on how to effectively measure and mitigate fairness in FL settings. One idea is to look into redefining the current notion of fairness altogether in ML, where fairness can become allowing equal access to accurate models to all participating clients. This, for example, can be examined through looking into personalized FL. However, personalization in FL can bring new tensions that may arose when studying the extremes of having a fair global model versus an accurate personalized local model.

Federated Learning in personalization: As user behavior gains traction in impacting future applications in industries such as healthcare, automotive, smart homes, and smart cities, personalization becomes a crucial feature. Personalization is based on customizing applications to fit user habits and preferences. The authors in Nadiger, Kumar, and Abdelhak (2019) present an approach of using Federated Reinforcement Learning (FRL) to speed up the personalization processes. The overall architecture of FRL utilizes a grouping policy, learning policy, and federation policy. The inherent setting of FL promotes a setting where local model updates are used to provide a more personalized experience to edge clients, in comparison to utilizing prediction from a global model alone. More work is expected on the front of utilizing FL for user personalization and studying the tension that potentially arises between personalization and performance.

Federated Learning in governments and public sectors: Numerous directions of high importance open up with considering FL under the scope of governmental and public sector applications. Future directions of FL within government services will look into growing the automation of government services, such as predictions of traffic flow and detection of public vigilance, without concerns of data privacy (Qiang et al., 2021). Advancements in FL under governmental governance are also crucial when looking towards achieving a smart or e-government infrastructure (Qiang et al., 2021). FL will allow private and secure data analysis across government departments and agencies without the need to expose each entity's unique data pool. The work of Balta et al. (2021) studies accountability of the FL processes and their implications on legislative and jurisdictional issues that generally arise in public sectors. Denny, Kazim, Koshiyama, Treleaven, and Dolga (2021) express the benefit that the private sector has seen over public in terms of technological supply-chain management and automation of processes. Their work (Denny et al., 2021) discusses interesting points about the importance of transparency and an open economy, to which we become closer by allowing secure access to data owned by public and government sectors. By achieving an open economy, it becomes possible to eliminate governmental corruption and secrecy to promote and foster a growing economy with trusted investment and innovation. The aforementioned bring forth a new set of considerations to be studied in future directions of FL within governments and public sectors.

7. Conclusion

Federated Learning (FL) is gaining a lot of attention recently from both academia and industry. Many new applications and use cases are seen utilizing FL to improve the quality and privacy of data-driven applications. This survey paints a big picture of FL technical fundamentals, considerations, and characterization that is suitable for those with a basic understanding of Machine Learning (ML). Furthermore, we examined the privacy factor in FL in detail under different aspects and considerations. More importantly, this survey puts a large emphasis on highlighting current trending applications and use cases that utilize FL in technology and markets. We scope applications of FL in the following technologies: Artificial Intelligence, Internet of Things, blockchain, Natural Language Processing, autonomous vehicles, and resource allocation, as well as in the following market sectors: data science, healthcare, education, governmental/public sectors, and industry. The survey also presents a summary of potential high-impact areas and questions to be studied under various FL applications within recommendation systems, autonomous vehicles, battery management, privacy, personalization, and government and public sectors. Through its comprehensive review, this survey acts a broad reference point for future explorations in the FL domain.

References

Aïvodji, U. M., Gambs, S., & Martin, A. (2019). IOTFLA : A secured and privacy-preserving smart home architecture implementing federated learning. In *2019 IEEE security and privacy workshops (SPW)* (pp. 175–180).

Al-Laith, A., Shahbaz, M., Alaskar, H. F., & Rehmat, A. (2021). Arasencorpus: A semi-supervised approach for sentiment annotation of a large arabic text corpus. *Applied Sciences*, *11*(5), 2434.

Alawadi, S., Kebande, V. R., Dong, Y., Bugeja, J., Persson, J. A., & Olsson, C. M. (2021). A federated interactive learning iot-based health monitoring platform. In *European conference on advances in databases and information systems* (pp. 235–246). Springer.

Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020a). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, *8*, 140699–140725.

Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020b). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, *8*, 140699–140725. http://dx.doi.org/10.1109/ACCESS.2020.3013541.

Ang, F., Chen, L., Zhao, N., Chen, Y., Wang, W., & Yu, F. R. (2020). Robust federated learning with noisy communication. *IEEE Transactions on Communications*.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115.

Awan, S., Li, F., Luo, B., & Liu, M. (2019). Poster: A reliable and accountable privacy-preserving federated learning framework using the blockchain. In *CCS '19, Proceedings of the 2019 ACM SIGSAC conference on computer and communications security* (pp. 2561–2563). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3319535.3363256.

Bagheri, B., Rezapoor, M., & Lee, J. (2020). A unified data security framework for federated prognostics and health management in smart manufacturing. *Manufacturing Letters*.

Balta, D., Sellami, M., Kuhn, P., Schöpp, U., Buchinger, M., Baracaldo, N., et al. (2021). Accountable federated machine learning in government: Engineering and management insights. In *International conference on electronic participation* (pp. 125–138). Springer.

Google AI defeats human Go champion. (2017). URL https://www.bbc.com/news/technology-40042581.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828.

Bhagoji, A. N., Chakraborty, S., Mittal, P., & Calo, S. (2019). In K. Chaudhuri, & R. Salakhutdinov (Eds.), *Proceedings of machine learning research*: *vol. 97*, *Analyzing federated learning through an adversarial lens* (pp. 634–643). Long Beach, California, USA: PMLR, URL http://proceedings.mlr.press/v97/bhagoji19a.html.

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning, Vol. 4*. Springer.

Blanquer, I., Brasileiro, F., Brito, A., Calatrava, A., Carvalho, A., Fetzer, C., et al. (2020). Federated and secure cloud services for building medical image classifiers on an intercontinental infrastructure. *Future Generation Computer Systems*, *110*, 119–134.

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., et al. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316.

Brik, B., Ksentini, A., & Bouaziz, M. (2020). Federated learning for UAVs-enabled wireless networks: Use cases, challenges, and open problems. *IEEE Access*, *8*, 53841–53849.

Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C., & Shi, W. (2018). Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics*, *112*, 59–67.

Cai, L., Lin, D., Zhang, J., & Yu, S. Dynamic sample selection for federated learning with heterogeneous data in fog computing.

Chai, Z., Fayyaz, H., Fayyaz, Z., Anwar, A., Zhou, Y., Baracaldo, N., et al. (2019). Towards taming the resource and data heterogeneity in federated learning. In *2019 USENIX conference on operational machine learning (OpML 19)* (pp. 19–21). Santa Clara, CA: USENIX Association, URL https://www.usenix.org/conference/opml19/presentation/chai.

Chan, W., Jaitly, N., Le, Q. V., & Vinyals, O. (2015). Listen, attend and spell. arXiv preprint arXiv:1508.01211.

Chen, Y.-T., Chuang, Y.-C., & Wu, A.-Y. A. (2020). Online extreme learning machine design for the application of federated learning. In *2020 2nd IEEE international conference on artificial intelligence circuits and systems (AICAS)* (pp. 188–192). IEEE.

Chen, H., Li, H., Xu, G., Zhang, Y., & Luo, X. (2020). Achieving privacy-preserving federated learning with irrelevant updates over E-health applications. In *ICC 2020 - 2020 IEEE international conference on communications (ICC)* (pp. 1–6). http://dx.doi.org/10.1109/ICC40277.2020.9149385.

Chen, Y., Luo, F., Li, T., Xiang, T., Liu, Z., & Li, J. (2020a). A training-integrity privacy-preserving federated learning scheme with trusted execution environment. *Information Sciences*, *522*, 69–79. http://dx.doi.org/10.1016/j.ins.2020.02.037, URL http://www.sciencedirect.com/science/article/pii/S0020025520301201.

Chen, Y., Luo, F., Li, T., Xiang, T., Liu, Z., & Li, J. (2020b). A training-integrity privacy-preserving federated learning scheme with trusted execution environment. *Information Sciences*, *522*, 69–79.

Chen, Y., Qin, X., Wang, J., Yu, C., & Gao, W. (2020). Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems*, 1.

Chen, Y., Sun, X., & Jin, Y. (2019). Communication-efficient federated deep learning with asynchronous model update and temporally weighted aggregation. arXiv preprint arXiv:1903.07424.

Chen, M., Yang, Z., Saad, W., Yin, C., Poor, H. V., & Cui, S. (2019). Performance optimization of federated learning over wireless networks. In *2019 IEEE global communications conference (GLOBECOM)* (pp. 1–6). IEEE.

Conway-Jones, D., Tuor, T., Wang, S., & Leung, K. K. (2019). Demonstration of federated learning in a resource-constrained networked environment. In *2019 IEEE international conference on smart computing (SMARTCOMP)* (pp. 484–486). IEEE.

Deist, T. M., Dankers, F. J., Ojha, P., Marshall], M. S., Janssen, T., Faivre-Finn, C., et al. (2020). Distributed learning on 20 000+ lung cancer patients – The personal health train. *Radiotherapy and Oncology*, *144*, 189–200.

Denny, D., Kazim, E., Koshiyama, A., Treleaven, P., & Dolga, R. (2021). Anticorruption techs to face a global economy-federated learning, open data catalogues, and "blockchain". Available at SSRN.

Dodge, S., & Karam, L. (2017). A study and comparison of human and deep learning recognition performance under visual distortions. In *2017 26th international conference on computer communication and networks (ICCCN)* (pp. 1–7). IEEE.

Doku, R., Rawat, D. B., & Liu, C. (2019). Towards federated learning approach to determine data relevance in big data. In *2019 IEEE 20th international conference on information reuse and integration for data science (IRI)* (pp. 184–192). IEEE.

Dolui, K., Cuba Gyllensten, I., Lowet, D., Michiels, S., Hallez, H., & Hughes, D. (2019). Towards privacy-preserving mobile applications with federated learning: The case of matrix factorization (poster). In *Proceedings of the 17th annual international conference on mobile systems, applications, and services* (pp. 624–625).

Dong, Y., Chen, X., Shen, L., & Wang, D. (2020). EaSTFLy: Efficient and secure ternary federated learning. *Computers & Security*, Article 101824.

Du, Z., Wu, C., Yoshinaga, T., Yau, K. A., Ji, Y., & Li, J. (2020). Federated learning for vehicular internet of things: Recent advances and open issues. *IEEE Open Journal of the Computer Society*, *1*, 45–61.

Duan, M., Liu, D., Chen, X., Tan, Y., Ren, J., Qiao, L., et al. (2019). Astraea: Self-balancing federated learning for improving classification accuracy of mobile deep learning applications. In *2019 IEEE 37th international conference on computer design (ICCD)* (pp. 246–254). IEEE.

Dwork, C. (2008). Differential privacy: A survey of results. In *International conference on theory and applications of models of computation* (pp. 1–19). Springer.

Elayan, H., Aloqaily, M., & Guizani, M. (2021a). Deep federated learning for IoT-based decentralized healthcare systems. In *2021 international wireless communications and mobile computing (IWCMC)* (pp. 105–109). IEEE.

Elayan, H., Aloqaily, M., & Guizani, M. (2021b). Digital twin for intelligent context-aware IoT healthcare systems. *IEEE Internet of Things Journal*.

Elayan, H., Aloqaily, M., & Guizani, M. (2021c). Sustainability of healthcare data analysis IoT-based systems using deep federated learning. *IEEE Internet of Things Journal*.

Encheva, S., & Tumin, S. (2008). On improving quality of the decision making process in a federated learning system. In *International conference on cooperative design, visualization and engineering* (pp. 192–195). Springer.

Fang, C., Guo, Y., Wang, N., & Ju, A. (2020). Highly efficient federated learning with strong privacy preservation in cloud computing. *Computers & Security*, Article 101889.

Feng, J., Rong, C., Sun, F., Guo, D., & Li, Y. (2020). Pmf: A privacy-preserving human mobility prediction framework via federated learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(1), 1–21.

Feraudo, A., Yadav, P., Safronov, V., Popescu, D. A., Mortier, R., Wang, S., et al. (2020). CoLearn: Enabling federated learning in MUD-compliant IoT edge networks. In *EdgeSys '20, Proceedings of the third ACM international workshop on edge systems, analytics and networking* (pp. 25–30). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3378679.3394528.

Franco, N., Van, H. M., Dreiser, M., & Weiss, G. (2021). Towards a self-adaptive architecture for federated learning of industrial automation systems. In *2021 international symposium on software engineering for adaptive and self-managing systems (SEAMS)*.

Gengler, N. (2019). Symposium review: Challenges and opportunities for evaluating and using the genetic potential of dairy cattle in the new era of sensor data from automation. *Journal of Dairy Science*, *102*(6), 5756–5763.

Gentry, C. (2009). Fully homomorphic encryption using ideal lattices. In *Proceedings of the forty-first annual ACM symposium on theory of computing* (pp. 169–178).

Ghosh, A. M., & Grolinger, K. (2021). Edge-cloud computing for internet of things data analytics: Embedding intelligence in the edge with deep learning. *IEEE Transactions on Industrial Informatics*.

Goecks, J., Jalili, V., Heiser, L. M., & Gray, J. W. (2020). How machine learning will transform biomedicine. *Cell*, *181*(1), 92–101.

Guo, Y., Wang, D., Vishwanath, A., Xu, C., & Li, Q. (2020). Towards federated learning for HVAC analytics: A measurement study. In *e-Energy '20*, *Proceedings of the eleventh ACM international conference on future energy systems* (pp. 68–73). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3396851.3397717.

Guo, S., Zeng, D., & Dong, S. (2020). Pedagogical data analysis via federated learning toward Education 4.0. *American Journal of Education and Information Technology*, *4*(2), 56.

Gururani, S., & Lerch, A. (2021). Semi-supervised audio classification with partially labeled data. In *2021 IEEE international symposium on multimedia (ISM)* (pp. 111–114). IEEE.

Han, X., Yu, H., & Gu, H. (2019). Visual inspection with federated learning. In *International conference on image analysis and recognition* (pp. 52–64). Springer.

Hänsch, R., & Hellwich, O. (2009). Semi-supervised learning for classification of polarimetric SAR-data. In *2009 IEEE international geoscience and remote sensing symposium, Vol. 3* (pp. III–987–III–990).

Hao, M., Li, H., Luo, X., Xu, G., Yang, H., & Liu, S. (2019). Efficient and privacy-enhanced federated learning for industrial artificial intelligence. *IEEE Transactions on Industrial Informatics*.

Hao, M., Li, H., Xu, G., Liu, S., & Yang, H. (2019). Towards efficient and privacy-preserving federated deep learning. In *ICC 2019 - 2019 IEEE international conference on communications (ICC)* (pp. 1–6).

Hard, A., Kiddon, C. M., Ramage, D., Beaufays, F., Eichner, H., Rao, K., et al. (2018). Federated learning for mobile keyboard prediction. URL https://arxiv.org/abs/1811.03604.

Hilmkil, A., Callh, S., Barbieri, M., Sütfeld, L. R., Zec, E. L., & Mogren, O. (2021). Scaling federated learning for fine-tuning of large language models. In *International conference on applications of natural language to information systems* (pp. 15–23).

Hsu, H.-Y., Srivastava, G., Wu, H.-T., & Chen, M.-Y. (2020). Remaining useful life prediction based on state assessment using edge computing on deep learning. *Computer Communications*, *160*.

Hu, R., Guo, Y., Li, H., Pei, Q., & Gong, Y. (2020). Personalized federated learning with differential privacy. *IEEE Internet of Things Journal*.

Huang, X., Ding, Y., Jiang, Z. L., Qi, S., Wang, X., & Liao, Q. (2020). DP-FL: a novel differentially private federated learning framework for the unbalanced data. *World Wide Web*, 1–17.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700–4708).

Huang, L., Shea, A. L., Qian, H., Masurkar, A., Deng, H., & Liu, D. (2019). Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *Journal of Biomedical Informatics*, *99*, Article 103291.

Imteaj, A., & Amini, M. H. (2019). Distributed sensing using smart end-user devices: Pathway to federated learning for autonomous IoT. In *2019 international conference on computational science and computational intelligence (CSCI)* (pp. 1156–1161).

Jalalirad, A., Scavuzzo, M., Capota, C., & Sprague, M. (2019). A simple and efficient federated recommender system. In *BDCAT '19*, *Proceedings of the 6th IEEE/ACM international conference on big data computing, applications and technologies* (pp. 53–58). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3365109.3368788.

James, B. T., Luczak, B. B., & Girgis, H. Z. (2018). MeShClust: an intelligent tool for clustering DNA sequences. *Nucleic Acids Research*, *46*(14), e83.

Jararweh, Y., Otoum, S., & Al Ridhawi, I. (2020). Trustworthy and sustainable smart city services at the edge. *Sustainable Cities and Society*, *62*, Article 102394.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, *14*(1–2), 1–210.

Kharitonov, E. (2019). *WSDM '19*, *Federated online learning to rank with evolution strategies* (pp. 249–257). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3289600.3290968.

Kim, Y. J., & Hong, C. S. (2019). Blockchain-based node-aware dynamic weighting methods for improving federated learning performance. In *2019 20th Asia-Pacific network operations and management symposium (APNOMS)* (pp. 1–4). IEEE.

Kim, H., Park, J., Bennis, M., & Kim, S.-L. (2019). Blockchained on-device federated learning. *IEEE Communications Letters*, *24*(6), 1279–1283.

Knight, W. (2020). Defeated chess champ garry kasparov has made peace with AI. URL https://www.wired.com/story/defeated-chess-champ-garry-kasparov-made-peace-ai/.

Konečnỳ, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527.

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. CoRR abs/1610.05492. arXiv:1610.05492. URL http://arxiv.org/abs/1610.05492.

Kuai, H., & Zhong, N. (2020). The extensible data-brain model: Architecture, applications and directions. *Journal of Computer Science*, Article 101103.

Kwon, D., Jeon, J., Park, S., Kim, J., & Cho, S. (2020). Multi-agent DDPG-based deep learning for smart ocean federated learning IoT networks. *IEEE Internet of Things Journal*, 1.

Lan, Y., Deng, H., & Chen, T. (2011). A new method of distance measure for graph-based semi-supervised learning. In *2011 international conference on machine learning and cybernetics, Vol. 4* (pp. 1444–1448).

Li, L., Fan, Y., Tse, M., & Lin, K.-Y. (2020). A review of applications in federated learning. *Computers & Industrial Engineering*, *149*, Article 106854. http://dx.doi.org/10.1016/j.cie.2020.106854, URL http://www.sciencedirect.com/science/article/pii/S0360835220305532.

Li, H., & Han, T. (2019). An end-to-end encrypted neural network for gradient updates transmission in federated learning. arXiv preprint arXiv:1908.08340.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50–60.

Li, Z., Sharma, V., & Mohanty, S. P. (2020). Preserving data privacy via federated learning: Challenges and solutions. *IEEE Consumer Electronics Magazine*, *9*(3), 8–16.

Li, L., Xiong, H., Guo, Z., Wang, J., & Xu, C.-Z. (2019). SmartPC: Hierarchical pace control in real-time federated learning system. In *2019 IEEE real-time systems symposium (RTSS)* (pp. 406–418). IEEE.

Lim, W. Y. B., Luong, N. C., Hoang, D. T., Jiao, Y., Liang, Y.-C., Yang, Q., et al. (2020). Federated learning in mobile edge networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, *22*(3), 2031–2063.

Lin, M. (2020). Accelerating the translation of artificial intelligence from ideas to routine clinical workflow. *Academic Radiology*, *27*(1), 121–122, Special Issue: Artificial Intelligence.

Lin, B. Y., He, C., Zeng, Z., Wang, H., Huang, Y., Soltanolkotabi, M., et al. (2021). FedNLP: A research platform for federated learning in natural language processing. arXiv preprint arXiv:2104.08815.

Lin, T., Kong, L., Stich, S. U., & Jaggi, M. (2020). Ensemble distillation for robust model fusion in federated learning. arXiv preprint arXiv:2006.07242.

Liu, W., Chen, L., Chen, Y., & Zhang, W. (2020). Accelerating federated learning via momentum gradient descent. *IEEE Transactions on Parallel and Distributed Systems*, *31*(8), 1754–1766.

Liu, L., Hu, Y., Yu, J., Zhang, F., Huang, G., Xiao, J., et al. (2019). Training encrypted models with privacy-preserved data on blockchain. In *Proceedings of the 3rd international conference on vision, image and signal processing* (pp. 1–6).

Liu, Y., Li, H., Xiao, J., & Jin, H. (2019). FLoc: Fingerprint-based indoor localization system under a federated learning updating framework. In *2019 15th international conference on mobile ad-hoc and sensor networks (MSN)* (pp. 113–118). IEEE.

Liu, X., Li, H., Xu, G., Lu, R., & He, M. (2020). Adaptive privacy-preserving federated learning. *Peer-to-Peer Networking and Applications*.

Liu, Y., Ma, Z., Yan, Z., Wang, Z., Liu, X., & Ma, J. (2020). Privacy-preserving federated k-means for proactive caching in next generation cellular networks. *Information Sciences*, *521*, 14–31.

Liu, B., Wang, L., & Liu, M. (2019). Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems. *IEEE Robotics and Automation Letters*, *4*(4), 4555–4562.

Liu, F., Wu, X., Ge, S., Fan, W., & Zou, Y. (2020). Federated learning for vision-and-language grounding problems. In *Proceedings of the AAAI conference on artificial intelligence, Vol. 34* (pp. 11572–11579). http://dx.doi.org/10.1609/aaai.v34i07.6824, (07). URL https://ojs.aaai.org/index.php/AAAI/article/view/6824.

Liu, L., Zhang, J., Song, S., & Letaief, K. B. (2019). Client-edge-cloud hierarchical federated learning. arXiv preprint arXiv:1905.06641.

Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2019). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. *IEEE Transactions on Industrial Informatics*, *16*(6), 4177–4186.

Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020). Federated learning for data privacy preservation in vehicular cyber-physical systems. *IEEE Network*, *34*(3), 50–56.

Lu, Y., Huang, X., Zhang, K., Maharjan, S., & Zhang, Y. (2020). Blockchain empowered asynchronous federated learning for secure data sharing in internet of vehicles. *IEEE Transactions on Vehicular Technology*, *69*(4), 4298–4311.

Lu, X., Liao, Y., Lió, P., & Hui, P. (2020). Privacy-preserving asynchronous federated learning mechanism for edge network computing. *IEEE Access*, *8*, 48970–48981.

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*, *29*(2), 102–127, Special Issue: Deep Learning in Medical Physics.

Luping, W., Wei, W., & Bo, L. (2019). Cmfl: Mitigating communication overhead for federated learning. In *2019 IEEE 39th international conference on distributed computing systems (ICDCS)* (pp. 954–964). IEEE.

Ma, C., Li, J., Ding, M., Yang, H. H., Shu, F., Quek, T. Q., et al. (2020). On safeguarding privacy and security in the framework of federated learning. *IEEE Network*.

Ma, X., Zhang, F., Chen, X., & Shen, J. (2018). Privacy preserving multi-party computation delegation for deep learning in cloud computing. *Information Sciences*, *459*, 103–116.

Majeed, U., & Hong, C. S. (2019). Flchain: Federated learning via MEC-enabled blockchain network. In *2019 20th Asia-Pacific network operations and management symposium (APNOMS)* (pp. 1–4). IEEE.

Malle, B., Giuliani, N., Kieseberg, P., & Holzinger, A. (2017). The more the merrier-federated learning from local sphere recommendations. In *International cross-domain conference for machine learning and knowledge extraction* (pp. 367–373). Springer.

Martinez, I., Francis, S., & Hafid, A. S. (2019). Record and reward federated learning contributions with blockchain. In *2019 international conference on cyber-enabled distributed computing and knowledge discovery (CyberC)* (pp. 50–57). IEEE.

Marulli, F., Bellini, E., & Marrone, S. (2020). A security-oriented architecture for federated learning in cloud environments. In *AINA workshops*.

McMahan, B., & Ramage, D. (2017). Federated learning: Collaborative machine learning without centralized training data. *Google AI Blog*.

McMahan, B., & Thakurta, A. (2021). Federated learning with formal differential privacy guarantees. *Google AI Blog*.

McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of bitcoin using machine learning. In *2018 26th Euromicro international conference on parallel, distributed and network-based processing (PDP)* (pp. 339–343). IEEE.

Mowla, N. I., Tran, N. H., Doh, I., & Chae, K. (2020). Federated learning-based cognitive detection of jamming attack in flying ad-hoc network. *IEEE Access*, *8*, 4338–4350.

Muniswamaiah, M., Agerwala, T., & Tappert, C. C. (2019). Federated query processing for big data in data science. In *2019 IEEE international conference on big data (Big data)* (pp. 6145–6147).

Nadiger, C., Kumar, A., & Abdelhak, S. (2019). Federated reinforcement learning for fast personalization. In *2019 IEEE second international conference on artificial intelligence and knowledge engineering (AIKE)* (pp. 123–127). IEEE.

Nguyen, D. C., Pathirana, P. N., Ding, M., & Seneviratne, A. (2019). Blockchain for 5g and beyond networks: A state of the art survey. arXiv preprint arXiv:1912.05062.

Nilsson, A., Smith, S., Ulm, G., Gustavsson, E., & Jirstrand, M. (2018). A performance evaluation of federated learning algorithms. In *DIDL '18*, *Proceedings of the second workshop on distributed infrastructures for deep learning* (pp. 1–8). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3286490.3286559.

Nishio, T., & Yonetani, R. (2019). Client selection for federated learning with heterogeneous resources in mobile edge. In *ICC 2019-2019 IEEE international conference on communications (ICC)* (pp. 1–7). IEEE.

Noble, B. D., Satyanarayanan, M., Narayanan, D., Tilton, J. E., Flinn, J., & Walker, K. R. (1997). Agile application-aware adaptation for mobility. *Operating Systems Review*, *31*(5), 276–287.

Nuding, F., & Mayer, R. (2020). Poisoning attacks in federated learning: An evaluation on traffic sign classification. In *Proceedings of the tenth ACM conference on data and application security and privacy* (pp. 168–170).

What is federated learning. (2020). https://blogs.nvidia.com/blog/2019/10/13/what-is-federated-learning/, Accessed: 2020-12-30.

Oneto, L., & Chiappa, S. (2020). Fairness in machine learning. In *Recent trends in learning from data* (pp. 155–196). Springer.

Pahwa, K., & Agarwal, N. (2019). Stock market analysis using supervised machine learning. In *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 197–200).

Pandey, S. R., Tran, N. H., Bennis, M., Tun, Y. K., Manzoor, A., & Hong, C. S. (2020). A crowdsourcing framework for on-device federated learning. *IEEE Transactions on Wireless Communication*, *19*(5), 3241–3256.

Pokhrel, S. R., & Choi, J. (2020a). Federated learning with blockchain for autonomous vehicles: Analysis and design challenges. *IEEE Transactions on Communications*.

Pokhrel, S. R., & Choi, J. (2020b). Improving TCP performance over WiFi for internet of vehicles: A federated learning approach. *IEEE Transactions on Vehicular Technology*, *69*(6), 6798–6802.

Posner, J., Tseng, L., Aloqaily, M., & Jararweh, Y. (2021). Federated learning in vehicular networks: opportunities and solutions. *IEEE Network*, *35*(2), 152–159.

Qian, Y., Hu, L., Chen, J., Guan, X., Hassan, M. M., & Alelaiwi, A. (2019). Privacy-aware service placement for mobile edge computing via federated learning. *Information Sciences*, *505*, 562–570.

Qiang, F., Lixin, T., Richard, L., et al. (2021). White paper-IEEE federated machine learning.

Qu, Y., Gao, L., Luan, T. H., Xiang, Y., Yu, S., Li, B., et al. (2020). Decentralized privacy using blockchain-enabled federated learning in fog computing. *IEEE Internet of Things Journal*.

Renuka, D. K., Hamsapriya, T., Chakkaravarthi, M. R., & Surya, P. L. (2011). Spam classification based on supervised learning using machine learning techniques. In *2011 international conference on process automation, control and computing* (pp. 1–7). IEEE.

Rey, V., Sánchez, P. M. S., Celdrán, A. H., & Bovet, G. (2022). Federated learning for malware detection in iot devices. *Computer Networks*, *204*, Article 108693.

Samarakoon, S., Bennis, M., Saad, W., & Debbah, M. (2018). Federated learning for ultra-reliable low-latency V2V communications. In *2018 IEEE global communications conference (GLOBECOM)* (pp. 1–7).

Samarakoon, S., Bennis, M., Saad, W., & Debbah, M. (2019). Distributed federated learning for ultra-reliable low-latency vehicular communications. *IEEE Transactions on Communications*, *68*(2), 1146–1159.

Samarakoon, S., Bennis, M., Saad, W., & Debbah, M. (2020). Distributed federated learning for ultra-reliable low-latency vehicular communications. *IEEE Transactions on Communications*, *68*(2), 1146–1159.

Saputra, Y. M., Hoang, D. T., Nguyen, D. N., Dutkiewicz, E., Mueck, M. D., & Srikanteswara, S. (2019). Energy demand prediction with federated learning for electric vehicle networks. In *2019 IEEE global communications conference (GLOBECOM)* (pp. 1–6).

Sarikaya, Y., & Ercetin, O. (2020). Regulating workers in federated learning by yardstick competition. In *VALUETOOLS '20, Proceedings of the 13th EAI international conference on performance evaluation methodologies and tools* (pp. 150–155). New York, NY, USA: Association for Computing Machinery.

Sariyildiz, M. B., Cinbis, R. G., & Ayday, E. (2020). Key protected classification for collaborative learning. *Pattern Recognition*, Article 107327.

Sattler, F., Wiedemann, S., Müller, K.-R., & Samek, W. (2019). Robust and communication-efficient federated learning from non-iid data. *IEEE Transactions on Neural Networks and Learning Systems*.

Savazzi, S., Nicoli, M., & Rampa, V. (2020). Federated learning with cooperating devices: A consensus approach for massive IoT networks. *IEEE Internet of Things Journal*, *7*(5), 4641–4654.

Savazzi, S., Nicoli, M., Rampa, V., & Kianoush, S. (2020). Federated learning with mutually cooperating devices: A consensus approach towards server-less model optimization. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 3937–3941). IEEE.

Shan, N., Cui, X., & Gao, Z. (2020). "DRL+ FL": An intelligent resource allocation model based on deep reinforcement learning for mobile edge computing. *Computer Communications*.

Sharma, P. K., Park, J. H., & Cho, K. (2020). Blockchain and federated learning-based distributed computing defence framework for sustainable society. *Sustainable Cities and Society*, Article 102220.

Shi, W., Zhou, S., & Niu, Z. (2019). Device scheduling with fast convergence for wireless federated learning. arXiv preprint arXiv:1911.00856.

Silva, S., Gutman, B. A., Romero, E., Thompson, P. M., Altmann, A., & Lorenzi, M. (2019). Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data. In *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)* (pp. 270–274).

Singh, J., Mittal, M., & Pareek, S. (2016). Customer behavior prediction using K-means clustering algorithm. In *Optimal inventory control and management techniques* (pp. 256–267). IGI Global.

Singh, S. K., Rathore, S., & Park, J. H. (2019). Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence. *Future Generation Computer Systems*.

Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., et al. (2020). Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *Advances in Neural Information Processing Systems*, *33*, 596–608.

Song, T., Tong, Y., & Wei, S. (2019). Profit allocation for federated learning. In *2019 IEEE international conference on big data (Big data)* (pp. 2577–2586). IEEE.

Sozinov, K., Vlassov, V., & Girdzijauskas, S. (2018). Human activity recognition using federated learning. In *2018 IEEE intl conf on parallel & distributed processing with applications, ubiquitous computing & communications, big data & cloud computing, social computing & networking, sustainable computing & communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)* (pp. 1103–1111). IEEE.

Sprague, M. R., Jalalirad, A., Scavuzzo, M., Capota, C., Neun, M., Do, L., et al. (2018). Asynchronous federated learning for geospatial applications. In *Joint European conference on machine learning and knowledge discovery in databases* (pp. 21–28). Springer.

Strangman, G. E., Sawyer, A., Fabre, K. M., Urquieta, E., Hury, J., Wu, J., et al. (2019). Deep-space applications for point-of-care technologies. *Current Opinion in Biomedical Engineering*, *11*, 45–50, Biomechanics and Mechanobiology: multiscale modeling • Novel Biomedical Technologies: Medical devices > point of care (LMIC).

Subramanyan, P., Sinha, R., Lebedev, I., Devadas, S., & Seshia, S. A. (2017). A formal foundation for secure remote execution of enclaves. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 2435–2450).

Sun, H., Li, S., Yu, F. R., Qi, Q., Wang, J., & Liao, J. (2020). Towards communication-efficient federated learning in the internet of things with edge computing. *IEEE Internet of Things Journal*, 1.

Sun, Y., Zhou, S., & Gündüz, D. (2019). Energy-aware analog aggregation for federated learning with redundant data. arXiv preprint arXiv:1911.00188.

Taïk, A., & Cherkaoui, S. Electrical load forecasting using edge computing and federated learning.

Tang, S., Zhou, W., Chen, L., Lai, L., Xia, J., & Fan, L. (2021). Battery-constrained federated edge learning in UAV-enabled IoT for B5G/6G networks. *Physical Communication*, *47*, Article 101381.

Toyoda, K., & Zhang, A. N. (2019). Mechanism design for an incentive-aware blockchain-enabled federated learning platform. In *2019 IEEE international conference on big data (Big data)* (pp. 395–403). IEEE.

Tran, N. H., Bao, W., Zomaya, A., Nguyen, M. N. H., & Hong, C. S. (2019). Federated learning over wireless networks: Optimization model design and analysis. In *IEEE INFOCOM 2019 - IEEE conference on computer communications* (pp. 1387–1395).

Tran, H.-V., Kaddoum, G., Elgala, H., Abou-Rjeily, C., & Kaushal, H. (2020). Lightwave power transfer for federated learning-based wireless networks. *IEEE Communications Letters*.

Triastcyn, A., & Faltings, B. (2019). Federated learning with Bayesian differential privacy. In *2019 IEEE international conference on big data (Big data)* (pp. 2587–2596). IEEE.

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R., et al. (2019). A hybrid approach to privacy-preserving federated learning. In *Proceedings of the 12th ACM workshop on artificial intelligence and security* (pp. 1–11).

Truex, S., Liu, L., Chow, K.-H., Gursoy, M. E., & Wei, W. (2020). LDP-Fed: federated learning with local differential privacy. In *Proceedings of the third ACM international workshop on edge systems, analytics and networking* (pp. 61–66).

Tseng, L., Yao, X., Otoum, S., Aloqaily, M., & Jararweh, Y. (2020). Blockchain-based database in an IoT environment: challenges, opportunities, and analysis. *Cluster Computing*, *23*(3), 2151–2165.

van Berlo, B., Saeed, A., & Ozcelebi, T. (2020). Towards federated unsupervised representation learning. In *EdgeSys '20, Proceedings of the third ACM international workshop on edge systems, analytics and networking* (pp. 31–36). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3378679.3394530.

Verma, D., White, G., & de Mel, G. (2019). Federated AI for the enterprise: A web services based implementation. In *2019 IEEE international conference on web services (ICWS)* (pp. 20–27). IEEE.

Wang, X., Han, Y., Wang, C., Zhao, Q., Chen, X., & Chen, M. (2019). In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Network*, *33*(5), 156–165.

Wang, Z., Song, M., Zhang, Z., Song, Y., Wang, Q., & Qi, H. (2019). Beyond inferring class representatives: User-level privacy leakage from federated learning. In *IEEE INFOCOM 2019 - IEEE conference on computer communications* (pp. 2512–2520).

Wang, Y., Tong, Y., & Shi, D. (2020). Federated latent Dirichlet allocation: A local differential privacy based framework. In *AAAI* (pp. 6283–6290).

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., et al. (2019). Adaptive federated learning in resource constrained edge computing systems. *IEEE Journal on Selected Areas in Communications*, *37*(6), 1205–1221.

Wang, F., Zhang, M., Wang, X., Ma, X., & Liu, J. (2020). Deep learning for edge computing applications: A state-of-the-art survey. *IEEE Access*, *8*, 58322–58336.

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., et al. (2020). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, *15*, 3454–3469.

Wu, Q., He, K., & Chen, X. (2020). Personalized federated learning for intelligent IoT applications: A cloud-edge based framework. *IEEE Open Journal of the Computer Society*, *1*, 35–44.

Xu, R., Baracaldo, N., Zhou, Y., Anwar, A., & Ludwig, H. (2019). Hybridalpha: An efficient approach for privacy-preserving federated learning. In *Proceedings of the 12th ACM workshop on artificial intelligence and security* (pp. 13–23).

Xu, G., Li, H., Liu, S., Yang, K., & Lin, X. (2020). VerifyNet: Secure and verifiable federated learning. *IEEE Transactions on Information Forensics and Security*, *15*, 911–926.

Xu, Z., Li, L., & Zou, W. (2019a). Exploring federated learning on battery-powered devices. In *Proceedings of the ACM turing celebration conference-China* (pp. 1–6).

Xu, Z., Li, L., & Zou, W. (2019b). Exploring federated learning on battery-powered devices. In *ACM TURC '19*, *Proceedings of the ACM turing celebration conference - China*. New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3321408.3323080.

Yan, M., Chen, B., Feng, G., & Qin, S. (2020). Federated cooperation and augmentation for power allocation in decentralized wireless networks. *IEEE Access*, *8*, 48088–48100.

Yang, Z., Chen, M., Saad, W., Hong, C. S., & Shikh-Bahaei, M. (2019). Energy efficient federated learning over wireless communication networks.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019a). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(2), 1–19.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019b). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, *10*(2).

Yang, H. H., Liu, Z., Quek, T. Q., & Poor, H. V. (2019). Scheduling policies for federated learning in wireless networks. *IEEE Transactions on Communications*, *68*(1), 317–333.

Yang, W., Zhang, Y., Ye, K., Li, L., & Xu, C.-Z. (2019). FFD: A federated learning based method for credit card fraud detection. In *International conference on big data* (pp. 18–32). Springer.

Yao, A. C.-C. (1986). How to generate and exchange secrets. In *27th annual symposium on foundations of computer science (Sfcs 1986)* (pp. 162–167). IEEE.

Yao, X., Huang, C., & Sun, L. (2018). Two-stream federated learning: Reduce the communication costs. In *2018 IEEE visual communications and image processing (VCIP)* (pp. 1–4).

Yao, X., Huang, T., Wu, C., Zhang, R., & Sun, L. (2019). Towards faster and better federated learning: A feature fusion approach. In *2019 IEEE international conference on image processing (ICIP)* (pp. 175–179). IEEE.

Yu, H., Liu, Z., Liu, Y., Chen, T., Cong, M., Weng, X., et al. (2020). A sustainable incentive scheme for federated learning. *IEEE Intelligent Systems*.

Zellinger, W., Wieser, V., Kumar, M., Brunner, D., Shepeleva, N., Gálvez, R., et al. (2021). Beyond federated learning: On confidentiality-critical machine learning applications in industry. *Procedia Computer Science*, 734–743.

Zeng, T., Semiari, O., Mozaffari, M., Chen, M., Saad, W., & Bennis, M. (2020). Federated learning in the sky: Joint power allocation and scheduling with UAV swarms. arXiv:2002.08196.

Zhang, X., Fu, A., Wang, H., Zhou, C., & Chen, Z. A Privacy-Preserving and Verifiable Federated Learning Scheme.

Zhang, J., Wang, J., Zhao, Y., & Chen, B. (2019). An efficient federated learning scheme with differential privacy in mobile edge computing. In *International conference on machine learning and intelligent communications* (pp. 538–550). Springer.

Zhao, Y., Chen, J., Wu, D., Teng, J., & Yu, S. (2019). Multi-task network anomaly detection using federated learning. In *SoICT 2019*, *Proceedings of the tenth international symposium on information and communication technology* (pp. 273–279). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3368926.3369705.

Zhao, Y., Chen, J., Zhang, J., Wu, D., Teng, J., & Yu, S. (2019). PDGAN: A novel poisoning defense method in federated learning using generative adversarial network. In *International conference on algorithms and architectures for parallel processing* (pp. 595–609). Springer.

Zhou, W., Li, Y., Chen, S., & Ding, B. (2018). Real-time data processing architecture for multi-robots based on differential federated learning. In *2018 IEEE smartworld, ubiquitous intelligence computing, advanced trusted computing, scalable computing communications, cloud big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (pp. 462–471).

Zhou, Z., Yang, S., Pu, L. J., & Yu, S. (2020). CEFL: Online admission control, data scheduling and accuracy tuning for cost-efficient federated learning across edge nodes. *IEEE Internet of Things Journal*.

Zhu, X., Wang, J., Hong, Z., Xia, T., & Xiao, J. (2019). Federated learning of unsegmented Chinese text recognition model. In *2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI)* (pp. 1341–1345). IEEE.

Ziller, A., Trask, A., Lopardo, A., Szymkow, B., Wagner, B., Bluemke, E., et al. (2021). Pysyft: A library for easy federated learning. In *Federated learning systems* (pp. 111–139). Springer.

Zou, Y., Feng, S., Niyato, D., Jiao, Y., Gong, S., & Cheng, W. (2019). Mobile device training strategies in federated learning: An evolutionary game approach. In *2019 international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData)* (pp. 874–879).

Zou, Y., Feng, S., Xu, J., Gong, S., Niyato, D., & Cheng, W. (2019). Dynamic games in federated learning training service market. In *2019 IEEE pacific rim conference on communications, computers and signal processing (PACRIM)* (pp. 1–6). IEEE.